ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ
ΣΧΟΛΗ ΤΕΧΝΟΛΟΓΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΜΗΧΑΝΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΒΙΟΜΗΧΑΝΙΑΣ
ΤΟΜΕΑΣ ΟΡΓΑΝΩΣΗΣ ΠΑΡΑΓΩΓΗΣ ΚΑΙ ΒΙΟΜΗΧΑΝΙΚΗΣ ΔΙΟΙΚΗΣΗΣ

ΒΕΛΤΙΣΤΟΠΟΙΗΣΗ ΠΡΟΒΛΕΨΗΣ ΑΞΙΟΠΙΣΤΙΑΣ ΜΕ ΤΗΝ ΧΡΗΣΗ ΟΜΑΔΟΠΟΙΗΜΕΝΩΝ ΜΕΤΑΒΛΗΤΩΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΑΛΕΞΗΣ - ΕΡΒΙΝ Σ. ΣΤΟΥΡΜ

ΕΠΙΒΛΕΠΩΝ: Δρ. Ι. Λ. ΜΠΑΚΟΥΡΟΣ

ΒΟΛΟΣ 1997

ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ
ΣΧΟΛΗ ΤΕΧΝΟΛΟΓΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΜΗΧΑΝΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΒΙΟΜΗΧΑΝΙΑΣ
ΤΟΜΕΑΣ ΟΡΓΑΝΩΣΗΣ ΠΑΡΑΓΩΓΗΣ ΚΑΙ ΒΙΟΜΗΧΑΝΙΚΗΣ ΔΙΟΙΚΗΣΗΣ

ΒΕΛΤΙΣΤΟΠΟΙΗΣΗ ΠΡΟΒΛΕΨΗΣ ΑΞΙΟΠΙΣΤΙΑΣ ΜΕ ΤΗΝ ΧΡΗΣΗ ΟΜΑΔΟΠΟΙΗΜΕΝΩΝ ΜΕΤΑΒΛΗΤΩΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΑΛΕΞΗΣ - ΕΡΒΙΝ Σ. ΣΤΟΥΡΜ

ΕΠΙΒΛΕΠΩΝ: Δρ. Ι. Λ. ΜΠΑΚΟΥΡΟΣ

ΒΟΛΟΣ 1997

στους γονείς μου

ΠΡΟΛΟΓΟΣ

Κίνητρο για την εκπόνηση της παρούσας διπλωματικής εργασίας αποτέλεσε η διδακτορική διατριβή του Ι. Λ. Μπακούρου. Στην διατριβή αυτή εφαρμόζονται πολυμεταβλητές στατιστικές τεχνικές για την μελέτη και πρόβλεψη της αξιοπιστίας υποθαλάσσιων αγωγών.

Η παρούσα μελέτη, με σκοπό την βελτιστοποίηση της πρόβλεψης αξιοπιστίας, εφαρμόζει δύο συγκεκριμένες πολυμεταβλητές στατιστικές τεχνικές κάνοντας, παράλληλα, χρήση ομαδοποιημένων μεταβλητών.

Τα κεφάλαια που ακολουθούν, έχουν χωριστεί σε δύο κύριες ενότητες. Στην πρώτη ενότητα, γίνεται μια αναφορά σε προηγούμενες μελέτες που σχετίζονται με την επίλυση προβλημάτων αξιοπιστίας βάσει πολυμεταβλητών στατιστικών τεχνικών. Επίσης, περιγράφονται και αναλύονται οι θεωρίες των δύο στατιστικών τεχνικών που χρησιμοποιούνται στην παρούσα μελέτη, της διαχωριστικής ανάλυσης και της ταξινόμησης. Τέλος, γίνεται μια σύντομη αναφορά στην σημασία της χρήσης και εφαρμογής ομαδοποιημένων μεταβλητών.

Στην δεύτερη ενότητα, ακολουθεί η αναλυτική παρουσίαση της πρακτικής εφαρμογής των τεχνικών της διαχωριστικής ανάλυσης και της ταξινόμησης με την χρήση ομαδοποιημένων μεταβλητών, σε πραγματικά δεδομένα υποθαλάσσιων αγωγών.

Στο σημείο αυτό, για την πραγματοποίηση και ολοκλήρωση της διπλωματικής αυτής εργασίας, θα ήθελα να απευθύνω τις ευχαριστίες μου σε όσους συνέβαλαν στον σκοπό αυτό, και συγκεκριμένα:

- τον Δρ. Χημικό Μηχανικό και Επίκουρο Καθηγητή κ. Ιωάννη Λ. Μπακούρο, η διδακτορική διατριβή του οποίου αποτέλεσε το θεωρητικό υπόβαθρο της

διπλωματικής αυτής εργασίας, για την καθοδήγηση που μου παρείχε κατά την συγγραφή του παρόντος τεύχους.

- τον Dr. De la Mare, Chairman of Technology Management Unit at the University of Bradford, ο οποίος εμπιστεύθηκε την παροχή των στοιχείων για τους υποθαλάσσιους αγωγούς μεταφοράς πετρελαίου, από την εταιρία Battelle της Ελβετίας, για εκπαιδευτικούς και μόνο σκοπούς.

Βόλος, Σεπτέμβριος 1997

ΑΛΕΞΗΣ - ΕΡΒΙΝ Σ. ΣΤΟΥΡΜ

ΠΕΡΙΕΧΟΜΕΝΑ

ΣΕΛΙΔΑ

ΕΝΟΤΗΤΑ Β : ΠΡΑΚΤΙΚΗ ΕΦΑΡΜΟΓΗ ΤΕΧΝΙΚΩΝ ΔΙΑΧΩΡΙΣΤΙΚΗΣ ΑΝΑΛΥΣΗΣ ΚΑΙ ΤΑΞΙΝΟΜΗΣΗΣ

Κεφάλαιο 6° : ΠΡΑΚΤΙΚΗ ΕΦΑΡΜΟΓΗ ΤΕΧΝΙΚΩΝ ΑΞΙΟΠΙΣΤΙΑΣ ΣΤΟΥΣ ΑΓΩΓΟΥΣ ΤΗΣ ΒΟΡΕΙΟΥ ΘΑΛΑΣΣΑΣ

ΕΙΣΑΓΩΓΗ

«Αξιοπιστία είναι η πιθανότητα ένα εξάρτημα, μια συσκευή ή ένα σύστημα να πραγματοποιήσει την προκαθορισμένη αποστολή του, χωρίς αποτυχίες, για ένα ορισμένο χρονικό διάστημα, όταν λειτουργεί σωστά σ' ένα καθορισμένο περιβάλλον».

Στον σχεδιασμό ενός μηχανικού τεχνουργήματος, συχνά ο σχεδιαστής αναγκάζεται να αλλάξει κάποια σχεδιαστικά χαρακτηριστικά έναντι κάποιων άλλων, στην ανάγκη αναζήτησης ενός συνόλου το οποίο θα αντεπεξέρχεται στις ανάγκες του πελάτη, κατά τον κοστολογικά αποτελεσματικότερο τρόπο. Στο παρελθόν, η αναζήτηση αυτή συχνά περιλάμβανε σχεδιαστικά κριτήρια σχετικά με χαρακτηριστικά όπως το βάρος, το μέγεθος, την ταχύτητα, την οικονομία καυσίμων καθώς και με αισθητικούς παράγοντες. Εντούτοις σήμερα, στην μελέτη και αναζήτηση σχεδίων και προτύπων που να ικανοποιούν τις απαιτήσεις ενός βέλτιστου κύκλου ζωής και να αντεπεξέρχονται στα θεσπισμένα standards ασφάλειας και υγιεινής, άλλα μεγέθη όπως η ποιότητα, η συντήρηση και η αξιοπιστία λαμβάνουν πλέον ιδιαίτερη σημασία.

Η αξιοπιστία αποτελεί ένα μέτρο εκτίμησης του βαθμού επιτυχίας ενός συστήματος. Έτσι, στην σύγχρονη διοικητική πρακτική πρέπει να είμαστε σε θέση

να προβλέπουμε την κατάσταση στην οποία θα βρίσκεται ένα μεμονωμένο μηχάνημα ή ένα ολόκληρο σύστημα, στο τέλος ενός καθορισμένου χρονικού ορίζοντα. Η θεωρία της αξιοπιστίας μας επιτρέπει κατά βάση τον προσδιορισμό του χρόνου λειτουργίας χωρίς το σύστημα να υποστεί βλάβη. Τα αποτελέσματα της εφαρμογής της θεωρίας αξιοπιστίας, σε συνδυασμό με τον σωστό προγραμματισμό συντήρησης και ένα δομημένο σύστημα λήψεως αποφάσεων, αναμφίβολα αποτελούν βασικά χαρακτηριστικά ενός σχεδίου με σκοπό την ελαχιστοποίηση των βλαβών του συστήματος και την κοστολογικά ορθολογικότερη προσέγγισή του.

Η κατάσταση στην οποία θα βρεθεί ένα σύστημα στο τέλος μια καθορισμένης περιόδου, εξαρτάται σε υψηλό βαθμό από τα ιδιαίτερα χαρακτηριστικά του. Η προσέγγιση του προβλήματος απαιτεί, επομένως, την μελέτη των χαρακτηριστικών αυτών. Η επίλυση στην περίπτωση αυτή γίνεται βάσει της στατιστικής θεωρίας πολυμεταβλητών διαδικασιών. Τέτοιες διαδικασίες είναι η διαχωριστική ανάλυση και η ταξινόμηση. Οι τεχνικές αυτές βρίσκουν εφαρμογή ευρέως σε διάφορους επιστημονικούς τομείς, όπως στις πολιτικές επιστήμες, στην ιατρική, στην έρευνα αγοράς κλπ., σαν μέσο διάγνωσης και πρόγνωσης.

Οι τεχνικές της διαχωριστικής ανάλυσης και της ταξινόμησης χρησιμοποιούνται και σε προβλήματα που απασχολούν τους μηχανικούς, όπως π.χ. σε θέματα αξιοπιστίας. Ο Sayles (1980) χρησιμοποίησε τεχνικές διαχωριστικών συναρτήσεων ως μέσο ταξινόμησης πληροφοριών αξιοπιστίας. Ο Betts (1984) χρησιμοποιώντας γραμμική διαχωριστική ανάλυση (Linear Discriminant Analysis) ήταν σε θέση να υπολογίσει την πιθανότητα αποτυχίας μιας εταιρίας, με υψηλό βαθμό ακρίβειας. Ο Ιωάννης Λ. Μπακούρος (1988) στην διδακτορική του διατριβή με τίτλο «Offshore Pipeline Reliability Prediction» μελέτησε την επίδραση ορισμένων χαρακτηριστικών (διάμετρος, πίεση, πάχος κ.ά.) στην αξιοπιστία υποθαλάσσιων πετρελαιαγωγών, κάνοντας χρήση πολυμεταβλητών στατιστικών τεχνικών.

Η παρούσα διπλωματική εργασία αποτελεί μια προσπάθεια συνέχειας της μελέτης αξιοπιστίας, στα πλαίσια του γενικότερου ερευνητικού έργου που έχει προηγηθεί. Κατ' επέκταση των εφαρμογών των τεχνικών της διαχωριστικής ανάλυσης και της ταξινόμησης, επιλέχθηκαν αντιπροσωπευτικά χαρακτηριστικά που σχετίζονται εμπειρικά με την αξιοπιστία των υποθαλάσσιων αγωγών με-

ταφοράς πετρελαίου, σε συνδυασμό όμως με την χρησιμοποίηση **ομαδοποιημένων χαρακτηριστικών**. Απώτερος σκοπός η αξιοποίηση κατά το βέλτιστο των ήδη υπαρχουσών μεταβλητών ώστε να διερευνηθεί η πρακτική σημασία της εφαρμογής της τεχνικής αυτής. Σημειώνεται στο σημείο αυτό ότι παρόμοια μελέτη, εισάγοντας την έννοια των ομαδοποιημένων μεταβλητών, έχει πραγματοποιηθεί μόνο μία φορά[1].

Οι τεχνικές της διαχωριστικής ανάλυσης και της ταξινόμησης, με την χρήση ομαδοποιημένων μεταβλητών βρίσκουν πρακτική εφαρμογή σε πραγματικά δεδομένα υποθαλάσσιων πετρελαιαγωγών που είναι εγκατεστημένοι στην Βόρειο Θάλασσα. Η αναλυτική παρουσίαση της μελέτης αυτής γίνεται στην δεύτερη ενότητα της παρούσας εργασίας.

Κρίνεται σκόπιμο στο σημείο αυτό, η συνοπτική παρουσίαση των βασικών αρχών των θεωριών της διαχωριστικής ανάλυσης και της ταξινόμησης, ώστε ο αναγνώστης να εξοικειωθεί με τις πιο σύνθετες έννοιες που ακολουθούν.

Η διαχωριστική ανάλυση και η ταξινόμηση είναι τεχνικές που επεξεργάζονται τις τιμές των χαρακτηριστικών που προέρχονται από έναν πληθυσμό ατόμων των οποίων η τελική κατάσταση είναι ήδη γνωστή. Έτσι, ο πληθυσμός μπορεί να χωριστεί σε ομάδες των οποίων τα άτομα έχουν την ίδια τελική κατάσταση. Στην συνέχεια μετρώντας τα χαρακτηριστικά ενός νέου ατόμου, προβλέπουν την τελική του κατάσταση, δηλαδή την ομάδα στην οποία θα ανήκει.

Η διαχωριστική ανάλυση ερευνά τις παρατηρημένες διαφορές των χαρακτηριστικών των ατόμων όταν οι αιτίες της ομοιότητας δεν είναι απόλυτα κατανοητές. Συγκεκριμένα, γνωρίζουμε τις τιμές των χαρακτηριστικών του κάθε ατόμου και την ομάδα στην οποία ανήκει, και χρησιμοποιούμε μια μαθηματική διαδικασία η οποία αναλύει τις διαφορές που παρουσιάζονται στις τιμές των χαρακτηριστικών των ατόμων, που ανήκουν σε διαφορετικές ομάδες.

Με την διαδικασία αυτή, είμαστε σε θέση να γνωρίζουμε την ομάδα στην οποία ανήκει ένα άτομο λαμβάνοντας υπόψη μόνο τις μετρούμενες τιμές των χαρακτηριστικών του. Κάτι τέτοιο δεν μπορεί να γίνει με μια πρόχειρη ματιά επειδή οι διαφορές σχεδόν ποτέ δεν είναι τόσο εμφανής ώστε να εντοπισθούν αμέσως. Ο σκοπός της διαχωριστικής ανάλυσης είναι να περιγράφει είτε γρα-

---

[1] «Βελτιστοποίηση πρόβλεψης αξιοπιστίας με την χρήση ομαδοποιημένων μεταβλητών», Διπλωματική εργασία Γ. Καπετανάκη, Α. Τσομίδου, Θεσσαλονίκη 1995.

φικά είτε αλγεβρικά αυτές τις διαφορές των ατόμων των ομάδων που εξετάζονται. Προσπαθεί να βρει «διαχωριστές» που να εκφράζουν τις διαφορές αυτές, και των οποίων οι αριθμητικές τιμές είναι τέτοιες ώστε οι ομάδες να διαχωρίζονται όσο το δυνατόν καλύτερα.

Η ταξινόμηση είναι μια διαδικασία η οποία στηρίζεται σε καλά θεμελιωμένες στατιστικές τεχνικές που μπορούν να χρησιμοποιηθούν για την κατάταξη ενός ατόμου σε κάποια από τις υπάρχουσες ομάδες, και του οποίου δεν είναι γνωστή η τελική κατάσταση. Η κατάταξη αυτή επιτυγχάνεται σύμφωνα με ορισμένες συνθήκες βελτιστοποίησης, όπως π.χ. ελαχιστοποιώντας το συνολικό κόστος λαθεμένης ταξινόμησης.

Γίνεται αντιληπτό λοιπόν, ότι οι παραπάνω τεχνικές μπορούν να εφαρμοστούν όταν ο αριθμός των ομάδων είναι τουλάχιστον ίσος με δύο. Στην συγκεκριμένη μελέτη, δημιουργούνται δύο ομάδες που ονομάζονται «επιτυχίες» και «αποτυχίες».

ΕΝΟΤΗΤΑ Α

ΔΙΑΧΩΡΙΣΤΙΚΗ ΑΝΑΛΥΣΗ

ΤΑΞΙΝΟΜΗΣΗ

ΟΜΑΔΟΠΟΙΗΜΕΝΕΣ ΜΕΤΑΒΛΗΤΕΣ

Κεφάλαιο

1

ΧΡΗΣΗ ΔΙΑΧΩΡΙΣΤΙΚΗΣ ΑΝΑΛΥΣΗΣ ΚΑΙ ΤΑΞΙΝΟΜΗΣΗΣ ΓΙΑ ΕΠΙΛΥΣΗ ΠΡΟΒΛΗΜΑΤΩΝ ΑΞΙΟΠΙΣΤΙΑΣ

1.1 Γενικά

Η αξιοπιστία είναι ένα μέτρο εκτίμησης του βαθμού επιτυχίας ενός συστήματος που έχει προκαθορισμένη διάρκεια αποστολής και προκαθορισμένες συνθήκες λειτουργίας. Η αξιοπιστία ενός συστήματος είναι έννοια σχετική. Κατ' αυτόν τον τρόπο η συμπεριφορά ενός συστήματος μπορεί να κρίνεται ικανοποιητική για μια εφαρμογή ενώ μη ικανοποιητική για μια άλλη με αυστηρότερες απαιτήσεις.

Η γνώση της αξιοπιστίας ενός συστήματος επιτρέπει τον σωστό προγραμματισμό της συντήρησής του, με αποτέλεσμα την μείωση βλαβών και κατά συνέπεια την ελαχιστοποίηση του αντίστοιχου κόστους. Η κατάσταση στην οποία θα βρεθεί ένα σύστημα όμως, εξαρτάται σε μεγάλο βαθμό από τα ιδιαίτερα χαρακτηριστικά του και επομένως μια προσπάθεια ανάλυσής τους θα αποτελούσε σημαντικό βοήθημα για την πρόβλεψη της συμπεριφοράς του συστήματος.

Η θεωρία της στατιστικής για πολυμεταβλητές διαδικασίες, παρέχει σημαντικά εργαλεία για την πραγματοποίηση της παραπάνω προσπάθειας. Τέτοιου είδους εργαλεία αποτελούν και οι τεχνικές της διαχωριστικής ανάλυσης και της ταξινόμησης.

1.2 Γενικότερες εφαρμογές των τεχνικών της διαχωριστικής ανάλυσης και ταξινόμησης

Οι τεχνικές του διαχωρισμού έχουν χρησιμοποιηθεί ευρέως από την επιστήμη της στατιστικής τα τελευταία πενήντα χρόνια. Ορισμένες από τις πρώτες εφαρμογές εμφανίστηκαν στις κοινωνικές επιστήμες και αφορούσαν βασικά ψυχολογικούς και επιμορφωτικούς τομείς (Tatsuoka, Tieteman). Εξάλλου οι Cox και Brandwood χρησιμοποίησαν την διαχωριστική ανάλυση για να τοποθετήσουν σε χρονολογική σειρά τα έργα του Πλάτωνα.

Οι πολιτικοί επιστήμονες ανακάλυψαν την χρησιμότητα των στατιστικών πολυμεταβλητών τεχνικών μελετώντας την συμπεριφορά των ψηφοφόρων (Klecka). Στην ιατρική οι τεχνικές αυτές προσφέρονται και χρησιμοποιούνται σαν ένα διαγνωστικό και προγνωστικό εργαλείο (Bulbrook, Armitage). Έχουν γίνει επίσης μελέτες στις οποίες η διαχωριστική ανάλυση χρησιμοποιείται στην έρευνα αγοράς.

Οι μηχανικοί που μελετούν προβλήματα αξιοπιστίας χρησιμοποίησαν επίσης τεχνικές διαχωρισμού. Ο Sayles (1980) χρησιμοποιούσε τεχνικές διαχωριστικών συναρτήσεων σαν μέσο ταξινόμησης πληροφοριών αξιοπιστίας. Κατέληξε ότι αυτές οι τεχνικές προσφέρουν μια μέθοδο ποσοτικοποίησης της σπουδαιότητας των παραμέτρων του περιβάλλοντος, σαν μέτρο της επιρροής τους επάνω σε μια συνάρτηση, οι οποίες στα περισσότερα προβλήματα αξιοπιστίας θα ασχολούνται με μία ή περισσότερες κατηγορίες αποτυχιών ή επιτυχιών. Επισημαίνεται επίσης ότι η διαδικασία είναι ιδανικά εφαρμόσιμη σε προβλήματα πολλών μεταβλητών και έχει, σαν αποτέλεσμα, μια σειρά παραμέτρων με κάποια βαρύτητα, οι οποίες ποσοτικοποιούν την σημασία κάθε μεταβλητής στο συνολικό αποτέλεσμα.

Επίσης ο Betts χρησιμοποιώντας γραμμική διαχωριστική ανάλυση ήταν ικανός να υπολογίσει την πιθανότητα αποτυχίας μιας εταιρίας με μεγάλο βαθμό ακρίβειας.

1.3 Γενικές αρχές διαχωριστικής ανάλυσης και ταξινόμησης

Η διαχωριστική ανάλυση και η ταξινόμηση είναι τεχνικές που επεξεργάζονται τις τιμές των χαρακτηριστικών, που προέρχονται από έναν πληθυσμό ατόμων των οποίων η τελική κατάσταση είναι ήδη γνωστή. Κατ' αυτόν τον τρόπο ο πληθυσμός χωρίζεται σε ομάδες των οποίων τα άτομα έχουν την ίδια τελική κατάσταση. Στην συνέχεια μετρώντας τα χαρακτηριστικά ενός ατόμου προβλέπουν την τελική του κατάσταση, δηλαδή την ομάδα στην οποία θα ανήκει.

Η διαχωριστική ανάλυση ερευνά τις διαφορές των χαρακτηριστικών των ατόμων όταν οι αιτίες της ομοιότητας δεν είναι απόλυτα κατανοητές. Έτσι, γνωρίζοντας τις τιμές των χαρακτηριστικών κάθε ατόμου και την ομάδα στην οποία ανήκει, καταλήγει σε μια μαθηματική διαδικασία, η οποία αναλύει τις διαφορές που παρουσιάζονται στις τιμές των χαρακτηριστικών των ατόμων, που ανήκουν σε διαφορετικές ομάδες.

Με την διαδικασία αυτή, εύκολα αναγνωρίζεται η ομάδα στην οποία ανήκει ένα άτομο λαμβάνοντας υπόψη μόνο τις μετρούμενες τιμές των χαρακτηριστικών του, κάτι που δεν μπορεί να γίνει με μια πρόχειρη ματιά, επειδή οι διαφορές σχεδόν πάντοτε δεν είναι τόσο εμφανείς ώστε να εντοπιστούν αμέσως. Η διαχωριστική ανάλυση είναι μια τεχνική που έχει ως σκοπό να περιγράψει είτε γραφικά είτε αλγεβρικά αυτές τις διαφορές των ατόμων των ομάδων που εξετάζονται. Επιδιώκει να βρει «διαχωριστές» που να εκφράζουν τις διαφορές αυτές, και των οποίων οι αριθμητικές τιμές είναι τέτοιες ώστε οι ομάδες να διαχωρίζονται όσο το δυνατόν καλύτερα.

Η ταξινόμηση είναι μια διαδικασία η οποία στηρίζεται σε καλά θεμελιωμένους κανόνες, που μπορούν να χρησιμοποιηθούν για την κατάταξη ενός ατόμου, του οποίου δεν είναι γνωστή η τελική κατάσταση, σε κάποια από τις προϋπάρχουσες ομάδες. Η κατάταξη ενός ατόμου επιτυγχάνεται σύμφωνα με ορισμένες συνθήκες βελτιστοποίησης, όπως ελαχιστοποιώντας το ποσοστό της λαθεμένης ταξινόμησης, ή πιο ρεαλιστικά, ελαχιστοποιώντας το συνολικό κόστος μιας όχι τόσο επιτυχημένης ταξινόμησης. Στην εφαρμογή που παρουσιάζεται στην συνέχεια της εργασίας ο πληθυσμός χωρίζεται σε δύο ομάδες, που ονομάζονται «επιτυχίες» και «αποτυχίες».

Τελικά αξιοποιούνται τα αποτελέσματα της διαχωριστικής ανάλυσης και της ταξινόμησης. Συγκεκριμένα, αναλύεται η σχέση που συνδέει τις δύο τεχνικές και παριστάνεται σε ένα διάγραμμα. Το διάγραμμα αυτό, είναι αρκετά χρήσιμο και πρακτικό αφού καθιστά δυνατή την πρόβλεψη, με εύκολο και γρήγορο τρόπο, της πιθανότητας ενός ατόμου να ανήκει στις επιτυχίες ή στις αποτυχίες βάσει των τιμών των χαρακτηριστικών του.

1.4 Η μελέτη αξιοπιστίας υποθαλάσσιων αγωγών με χρήση διαχωριστικής ανάλυσης και ταξινόμησης

Πέρα από τις μελέτες των Sayles και Betts, ένα σημαντικό επιστημονικό πεδίο, στο οποίο έχουν εφαρμοστεί οι πολυμεταβλητές στατιστικές τεχνικές της διαχωριστικής ανάλυσης και ταξινόμησης είναι αυτό της μελέτης της αξιοπιστίας των υποθαλάσσιων αγωγών μεταφοράς πετρελαίου και φυσικού αερίου. Στις επόμενες παραγράφους, περιγράφονται εν συντομία, οι λόγοι που ευνοούν τη χρήση αγωγών για την μεταφορά τους, καθώς και οι σημαντικότερες επιστημονικές προσεγγίσεις που έχουν έως τώρα δημοσιευτεί γύρω από το θέμα.

1.4.1 Η σπουδαιότητα της χρήσης υποθαλάσσιων αγωγών για την μεταφορά προϊόντων του πετρελαίου

Αγωγοί για την μεταφορά ρευστών έχουν χρησιμοποιηθεί εδώ και πολλούς αιώνες. Αγωγοί από bamboo χρησιμοποιήθηκαν για πρώτη φορά από τους Κινέζους πριν από χιλιάδες χρόνια. Σε επόμενους αιώνες, ξύλινοι αγωγοί χρησιμοποιούνταν στην Μεγάλη Βρετανία. Εντούτοις, μόνο με την έλευση του χυτοσιδήρου κατέστη δυνατή η κατασκευή αγωγών οι οποίοι θα μπορούσαν να λειτουργήσουν υπό συνθήκες υψηλής πίεσης. Χρησιμοποιήθηκαν εκτεταμένα κατά την διάρκεια του 19^{ου} αιώνα και χρησιμοποιούνται ακόμη και σήμερα. Αγωγοί από χάλυβα κάνουν την εμφάνισή τους κατά τα τέλη του τελευταίου αιώνα,

διευκολύνοντας έτσι την κατασκευή αγωγών μικρών και μεγάλων διαμέτρων. Η συνεχώς αυξανόμενη χρήση υψηλής ποιότητας χάλυβα και μεγάλων εργοστασίων έλασης μετάλλου, επέτρεψαν την κατασκευή πολύ μεγάλων και ανθεκτικών αγωγών, με διαμέτρους που ξεπερνούν τα 3 μέτρα υπό πίεση που υπερβαίνει τα 10 Newton ανά τετραγωνικό χιλιοστόμετρο. Επίσης, η εξέλιξη των τεχνικών συγκόλλησης επέτρεψε την κατασκευή διαμήκως και περιφερειακά ή ελικοειδώς συγκολλημένων αγωγών. Αγωγοί επίσης κατασκευάζονται από ενισχυμένο σκυρόδεμα, προενταμένο σκυρόδεμα, κονία αμιάντου, πλαστικά κλπ., ανάλογα την ιδιάζουσα περίπτωση. Επιπλέον, εξελίξεις στην τεχνολογία ακριβούς μελέτης και πρόβλεψης της ροής εντός των αγωγών, ώθησαν εξίσου την χρήση τους.

Κατά τον προηγούμενο αιώνα, νερό και λύματα ήταν πρακτικά τα μόνο ρευστά που μεταφέρονταν μέσω αγωγών, ενώ σήμερα αποτελούν τα πιο κοινά μέσα μεταφοράς αερίων και πετρελαίου κατά μήκος μεγάλων αποστάσεων. Χημικά ρευστά και στερεά υπό την μορφή υδαρούς κονίας ή σε containers μεταφέρονται επίσης με διαρκώς αυξανόμενους ρυθμούς, με αποτέλεσμα σήμερα οι εν λειτουργία αγωγοί παγκοσμίως να υπερβαίνουν τα δύο εκατομμύρια χιλιόμετρα.

Αναμφισβήτητα, η ενεργειακή κρίση, επηρεάζοντας την εξέλιξη στους χώρους άντλησης πετρελαίου ανοικτής θαλάσσης, είχε σημαντικές επιπτώσεις στην χρήση των αγωγών φυσικού αερίου και πετρελαίου. Αρχικά, το φόρτωμα πετρελαιοφόρων μέσω αγκυροβολίων τόξων δια βραχέων αγωγών από τις πλατφόρμες παραγωγής, ήταν ο προτιμώμενος τρόπος αξιοποίησης των χώρων άντλησης πετρελαίου, αλλά σύντομα αντικαταστάθηκαν από μεγάλου μήκους αγωγούς από τις πλατφόρμες παραγωγής ως την στεριά.

Τα ιδιαίτερα πλεονεκτήματα των αγωγών πετρελαίου είναι:

- ✓ Είναι συνήθως ο πιο οικονομικός τρόπος μεταφοράς αφού εξασφαλίζουν μικρότερη σε έκταση διαδρομή από την συνήθη διαδρομή των τάνκερ και πλεονεκτούν για λιμάνια τα οποία είναι υποχρεωμένα να δέχονται μόνο μικρά σε μέγεθος και άρα αντιοικονομικότερα τάνκερς.
- ✓ Το κόστος τους δεν είναι τόσο ευάλωτο σε «εξάρσεις» στις τιμές του πετρελαίου, αφού το κύριο μέρος του αποτελείται από τη δαπάνη κεφαλαίου,

ενώ το κόστος λειτουργίας τους είναι σχετικά χαμηλό. Επιπλέον περίπου το 50% των δαπανών λειτουργίας αποτελείται από σταθερά στοιχεία, όπως συντήρηση, φορολογία κλπ.

- ✓ Σε σύγκριση με άλλες μορφές μεταφοράς, οι αγωγοί έχουν το πλεονέκτημα ότι δεν υπόκεινται σε αύξηση του κόστους λειτουργίας τους κατά τις περιόδους εισόδου και εξόδου του εμπορεύματος. Αντίθετα τα τάνκερ υπόκεινται σε φορολογία κατά τον ελλιμενισμό τους.
- ✓ Οι λειτουργίες των αγωγών δεν είναι ευπαθείς σε ανθρώπινα λάθη. Αυτό γιατί τα περισσότερα μοντέρνα συστήματα λειτουργούν αυτόματα και δεν απαιτούν ιδιαίτερη προσοχή.
- ✓ Οι θαμμένοι αγωγοί στον βυθό της θάλασσας δεν αποτελούν εμπόδιο για την ναυσιπλοΐα.
- ✓ Λόγω του ότι είναι τοποθετημένοι κάτω από το έδαφος, δεν καταστρέφουν το φυσικό περιβάλλον.
- ✓ Είναι πολύ ασφαλείς απέναντι σε ενδεχόμενη δολιοφθορά, καθώς είναι θαμμένοι στον βυθό.
- ✓ Η αξιοπιστία των αγωγών είναι συνήθως πολύ μεγαλύτερη απ' όλα τα άλλα μέσα μεταφοράς.

Υπάρχουν φυσικά και μειονεκτήματα σχετιζόμενα με τα συστήματα των αγωγών πετρελαίου όπως:

- ✗ Το αρχικό κόστος κεφαλαίου είναι συνήθως υψηλό, οπότε αν δεν είμαστε βέβαιοι για την έκταση της ζήτησης θα πρέπει αυτή να υποτεθεί σε κάποιο βαθμό, ώστε να μπορέσει να καθοριστεί το μέγεθος των αγωγών.
- ✗ Αλλαγή της διαδρομής εγκατεστημένων ήδη αγωγών, δεν μπορεί να πραγματοποιηθεί χωρίς να προκαλέσει υπερβολικό κόστος. Αντίθετα, μπορεί να εξασφαλιστεί αύξηση της χωρητικότητας με σχετικά χαμηλό κόστος.
- ✗ Ενδέχεται να παρουσιαστούν ορισμένα προβλήματα σχετικά με την άντληση στερεών, υγρών και αερίων τα οποία είναι δυνατόν να οδηγήσουν σε δυσλειτουργία των αγωγών.
- ✗ Είναι συχνά δύσκολο να εντοπιστεί η πηγή μικρών διαρροών ή μπλοκαρισμάτων σε μεγάλου μήκους αγωγούς.

- ✗ Η συντήρηση και επισκευή αγωγών τοποθετημένων κάτω από το έδαφος είναι πάντα πολύ δύσκολη.
- ✗ Οι αγωγοί συνήθως δεν μπορούν να χρησιμοποιηθούν για την μεταφορά περισσοτέρων του ενός υλικών, παρότι υπάρχουν αγωγοί μεταφοράς πολλών προϊόντων που λειτουργούν σε ομαδική βάση.

1.4.2 Η ανάγκη για υψηλή αξιοπιστία των αγωγών πετρελαίου και οι ενδεχόμενες οικονομίες από την χρήση τους

Η αυξανόμενη ένταση των παράκτιων δραστηριοτήτων, παρακινούμενη από την πετρελαϊκή κρίση, είχε ως αποτέλεσμα την εγκατάσταση ενός σημαντικού αριθμού πετρελαιαγωγών μεγάλου μήκους που μεταφέρουν αργό πετρέλαιο και φυσικό αέριο, στοιχεία ζωτικής σημασίας για την οικονομία αρκετών χωρών. Το εξαιρετικά υψηλό κόστος κεφαλαίου τέτοιων αγωγών σε συνδυασμό με το τεράστιο κόστος ανάπτυξης των περιοχών ύπαρξης πετρελαίου, οδήγησε στην ανάγκη εστίασης της προσοχής στην αξιοπιστία των αγωγών αυτών, ειδικά καθώς σχετίζεται αυτή με τα χαμένα έσοδα ή το ευκαιριακό κόστος που προκύπτουν σε περιπτώσεις μειζόνων αστοχιών των αγωγών και το πολύ υψηλό κόστος συντήρησης που απαιτείται για την επίλυση τέτοιων προβλημάτων.

Η λειτουργία τέτοιων μεγάλων συστημάτων πετρελαιαγωγών καθίσταται επίσης ιδιαίτερα περίπλοκη καθώς πολλές περιοχές είναι συνδεδεμένες σε ένα κοινό δίκτυο, το οποίο απαιτεί την αξιόπιστη λειτουργία καθενός τμήματος του δικτύου. Συνεπώς, ένας παράγοντας αύξουσας σημασίας περιλαμβάνει τις οικονομικές εμπλοκές της αξιοπιστίας των αγωγών, δηλαδή η μηχανολογική πολυπλοκότητα ενός συστήματος αγωγών πετρελαίου πρέπει να μεταφρασθεί σε όρους οικονομικούς. Επιτακτική δηλαδή καθίσταται η ανάγκη προσαρμογής της παρούσας επιστημονικής γνώσης στα δεδομένα και τις συνθήκες ανοικτής θαλάσσης, καθώς και της αναγνώρισης των ιδιαίτερων διαφορών μεταξύ των λειτουργιών στην στεριά και στην θάλασσα, ειδικά όσον αφορά τις δυσκολίες και το υψηλό κόστος σχετιζόμενο με υποθαλάσσιες επιθεωρήσεις, συντήρηση και επισκευή.

Λόγω της εύλογης ανησυχίας των σχεδιαστών μηχανικών όσον αφορά το υψηλό κόστος ενδεχόμενων επισκευών ή οικολογικών καταστροφών ως αποτέλεσμα αστοχιών, ο σχεδιασμός των υποβρύχιων πετρελαιαγωγών γίνεται βάσει αυστηρών μηχανολογικών κριτηρίων ώστε να μπορούν να αντεπεξέλθουν στις πιο δυσμενής συνθήκες που είναι δυνατόν να επιβληθούν λόγω της δριμύτητας του περιβάλλοντος στο οποίο εγκαθίστανται. Η περιβαλλοντική μόλυνση και τα πολιτικά προβλήματα που ακολουθούν την αστοχία ενός υποθαλάσσιου πετρελαιαγωγού αποτελούν αρκετά σημαντικούς λόγους για την εξασφάλιση του υψηλότερου βαθμού αξιοπιστίας.

Μια μείζονα απαίτηση για την εκτίμηση της αξιοπιστίας ενός πετρελαιαγωγού αποτελούν τα ιστορικά δεδομένα που σχετίζονται με τις αποτυχίες κατά την λειτουργία παρομοίων αγωγών. Δυστυχώς, υφίσταται έλλειψη τέτοιων στοιχείων, παρόλο που πλήθος δεδομένων υπάρχουν για τους αγωγούς στην στεριά. Εντούτοις, η άμεση εφαρμογή τους στους αγωγούς ανοικτής θαλάσσης, πέρα ορισμένων ειδικών περιπτώσεων, είναι συχνά εσφαλμένη. Είναι προφανές επομένως, ότι η συλλογή σχετικών στοιχείων και η κατάλληλη ανάλυσή τους αποτελεί έναν παράγοντα κλειδί στον καθορισμό της αξιοπιστίας υποθαλάσσιων αγωγών. Παρόλη την διεύρυνση των δραστηριοτήτων ανοικτής θαλάσσης κατά την διάρκεια των τελευταίων τριών δεκαετιών, μόνο λίγοι ερευνητές και οργανισμοί έχουν δημοσιεύσει μελέτες που να σχετίζονται με την αξιοπιστία των αγωγών, με σκοπό την γνωστική εμβάθυνση στους ενυπάρχοντες ρυθμούς αποτυχιών, την κοστολογικά αποτελεσματική αξιοπιστία και την βελτίωση της ασφάλειας. Στα πλαίσια αυτά, οι μελέτες των Cannon (1985), Battelle (1983) και de la Mare (1980), είναι οι πιο σημαντικές.

1.4.3 Προηγούμενες Προσεγγίσεις

- *Προσέγγιση De la Mare.*

Η αναφορά των De la Mare και Anderson (1980) είχε σαν σκοπό να εκτιμήσει τα χαρακτηριστικά της αξιοπιστίας, στην συγκεκριμένη περίπτωση, για τους πετρελαιαγωγούς της Βορείου Θάλασσας. Παρόλη την έλλειψη διαθέσι-

μων δεδομένων αποτυχιών, το πεδίο της μελέτης αυτής επεκτάθηκε ώστε να συμπεριλάβει και άλλα μεγάλα συστήματα πετρελαιαγωγών ανοικτής θαλάσσης καθ' όλο τον κόσμο, όπως την περίπτωση του κόλπου του Μεξικού, με σκοπό να ερευνηθεί αν οι ρυθμοί αποτυχιών παρουσιάζονταν επαρκώς παρεμφερείς με αυτούς των αγωγών της Βορείου Θάλασσας. Ο σκοπός της μελέτης αυτής περιορίστηκε στις γραμμές μεταφοράς αργού πετρελαίου, φυσικού αερίου και άλλων προϊόντων πετρελαίου. Η αποτυχία μιας γραμμής μεταφοράς καθοριζόταν ως κάθε γεγονός που είχε ως αποτέλεσμα την διαρροή του προϊόντος στην γραμμή αυτή. Διαρροές από αγωγούς συγκέντρωσης, σταθμούς άντλησης, βαλβίδες, συστήματα οργάνων καθώς και κάθε ζημιά σε κάποια γραμμή μεταφοράς που δεν είχε ως συνέπεια διαρροή αποκλείσθηκαν από την μελέτη αυτή, γιατί δεν αποτελούσαν αστοχίες αγωγών όπως αυτές καθορίζονται εδώ. Στην μελέτη, οι ταξινόμηση των αποτυχιών γίνεται βάσει τεσσάρων θεμελιωδών κατηγοριών αστοχιών:

1. **Κατηγορία A** (Μηχανική)	Λόγω κατασκευαστικών ατελειών-στρέβλωση, κόπωση, ελαττωμάτων υλικού και σφάλματα συγκολλήσεων.
2. **Κατηγορία C** (Διάβρωση)	Λόγω εσωτερικής και εξωτερικής διάβρωσης.
3. **Κατηγορία D** (Φυσικοί κίνδυνοι)	Λόγω ασταθούς και ολισθηρού εδάφους, καταιγίδων, αέρα και πλημμύρων.
4. **Κατηγορία E** (Τρίτοι παράγοντες)	Λόγω σύγκρουσης εκσκαφέων, αγκύρων κλπ.
5. **Λοιπά**	Αταξινόμητα αίτια.

Στην μελέτη αυτή λήφθηκε υπόψη ότι η συμπεριφορά των αποτυχιών των αγωγών, δεδομένου ότι οι αγωγοί θεωρούνται ως ισοδύναμες οντότητες, μπορεί να μην είναι η κατάλληλη, αφού όσο μεγαλύτερο είναι το μήκος του αγωγού τόσο μεγαλύτερη είναι και η πιθανότητα αποτυχίας, σε ένα καθορισμένο χρονικό διάστημα.

Η εκτιμήτρια του ρυθμού αποτυχιών *R*, η οποία χρησιμοποιείται συνήθως στην τεχνολογία των αγωγών, έχει καθοριστεί από την ακόλουθη μαθηματική έκφραση:

$$R = \frac{N}{T \cdot L_i},$$

όπου N: ο αριθμός των αποτυχιών των αγωγών που λαμβάνουν χώρα σε κάποιο καθορισμένο χρονικό διάστημα,

L_i: το μήκος του αγωγού που λειτουργεί σ' αυτό το χρονικό διάστημα,

T: ο συνολικός αριθμός αγωγών που λειτουργούν σ' αυτό το χρονικό διάστημα.

➪ *Προσέγγιση Det Norske Veritas.*

Η αναφορά από την Det Norske Veritas (1980) έδωσε μια γενική εικόνα των αστοχιών των αγωγών στις περιοχές της Βορείου Θάλασσας και του Κόλπου του Μεξικού. Ως αποτυχία ενός αγωγού θεωρείτο κάθε συμβάν το οποίο οδηγούσε σε σημαντικές διαρροές ή διαφορετικά απαιτούσε άμεση επισκευή.

Παρά το γεγονός ότι πολλοί από τους αγωγούς της Βορείου Θάλασσας βρίσκονταν σε λειτουργία για λιγότερο από δύο χρόνια, και επομένως η λειτουργική εμπειρία ήταν μάλλον περιορισμένη, παρουσιάστηκε μια λίστα ζημιών τις οποίες είχαν υποστεί οι αγωγοί μετά την τοποθέτησή τους αλλά και πριν τεθούν σε λειτουργία, καθώς επίσης και μια λίστα ζημιών που υπέστησαν οι αγωγοί κατά την διάρκεια της λειτουργίας τους.

Τα βασικά συμπεράσματα που προέκυψαν από την ανάλυση αυτή ήταν:

a) Κανένας ρυθμός αποτυχιών δεν δόθηκε λόγω της περιορισμένης λειτουργικής εμπειρίας,
b) Οι περισσότερες αποτυχίες παρουσιάστηκαν κατά το στάδιο της εκσκαφής και τοποθέτησης των αγωγών,
c) Οι πλειοψηφία των «λειτουργικών» αποτυχιών συνέβησαν σε ρηχά νερά ή κοντά σε πλατφόρμες.

Για την περίπτωση του Κόλπου του Μεξικού, ο φάκελος κάλυπτε αποτυχίες που είχαν ως συνέπεια ρήγματα ή διαρροή αργού πετρελαίου και φυσικού αερίου. Ένα σύνολο 136 αποτυχιών αναφέρθηκαν κατά την περίοδο 1967-1977, εκ των οποίων 94 έλαβαν χώρα σε αγωγούς διαμέτρου μικρότερης των 10 ιντσών, και 42 σε αγωγούς διαμέτρου μεγαλύτερης ή ίσης των 10 ιντσών. Εκτιμήθηκε ότι η πλειονότητα των αγωγών μικρής διαμέτρου ήταν ενδοδικτυακές γραμμές μεταφοράς, διφασικής ροής. Έτσι, ο ερευνητής θεώρησε ότι οι αγωγοί αυτοί δεν ήταν αντιπροσωπευτικοί αγωγών μεταφοράς και επομένως δεν συμπεριλήφθησαν στην περαιτέρω μελέτη.

Τα κύρια συμπεράσματα από την έρευνα αυτή ήταν:

a) Οι αγωγοί πετρελαίου παρουσίαζαν πολύ περισσότερες αποτυχίες από τους αγωγούς αερίου,
b) Δραστηριότητες σε περιοχές ανοικτής θαλάσσης (αγκυροβόληση κλπ.) και φαινόμενα εσωτερικής διάβρωσης ήταν οι κύριες ομάδες αιτιών των αποτυχιών,
c) Ο ρυθμός αποτυχιών μειωνόταν με την αύξηση της διαμέτρου των αγωγών,
d) Για αγωγούς διαμέτρου μεγαλύτερης των 20 ιντσών (εκ των οποίων το 95% περίπου ήταν αγωγοί αερίου), το πλήθος των αποτυχιών και ο αντίστοιχος ρυθμός βλαβών ήταν αισθητά μικρότερος σε σχέση με αυτόν των αγωγών διαμέτρου κάτω των 20 ιντσών.

➪ *Προσέγγιση Veritec*

Η μελέτη του Veritec (1984) ασχολήθηκε με ορισμένους σημαντικούς παράγοντες οι οποίοι είχαν ως σκοπό να εξασφαλίσουν την χρήση αγωγών υψηλής χωρητικότητας αλλά και την διαθεσιμότητα μεγάλου μήκους αγωγών με υψηλή λειτουργικότητα. Το έργο επικεντρώθηκε στα παρακάτω σημεία:

α) οικονομικοί παράγοντες
β) έννοια της διαθεσιμότητας
γ) κατηγορίες αποτυχιών των αγωγών και ρυθμοί των αποτυχιών

δ) σχεδιαστικές μελέτες.

Μια γενική εικόνα των αιτιών και των συχνοτήτων των αποτυχιών υποβρυχίων και μη αγωγών, οι οποίοι μεταφέρουν αργό πετρέλαιο και φυσικό αέριο, αποτέλεσε μέρος της μελέτης αυτής. Επίσης, διεξήχθη μια ανάλυση των αιτίων των αποτυχιών και ερευνήθηκε εν συντομία η επίδραση διαφόρων παραγόντων όπως η ηλικία του αγωγού, η τοποθεσία, η διάμετρος και το μεταφερόμενο προϊόν, σε σχέση με τον ρυθμό αποτυχιών.

Στην έρευνα αυτή, δόθηκαν οι ακόλουθοι ορισμοί:

Αποτυχία αγωγού:	Κάθε τυχαίο γεγονός ή κακή λειτουργία η οποία απαιτεί επισκευή και προκαλεί διακοπή λειτουργίας του αγωγού.
Ρυθμός αποτυχιών:	Η συχνότητα των αποτυχιών των αγωγών ανά χρόνο και μονάδα μήκους.
Λειτουργική εμπειρία:	Μήκος του αγωγού και χρόνια λειτουργίας.

Οι αποτυχίες των αγωγών ταξινομήθηκαν στις παρακάτω κατηγορίες:

- Μηχανικές αποτυχίες (δευτερεύων εξοπλισμός όπως βαλβίδες, σύνδεσμοι, φλάντζες, ως αποτέλεσμα κακής λειτουργίας που απαιτούν επισκευή).
- Αποτυχία γραμμής (ελαττωματικά υλικά, διάβρωση, απουσία προστατευτικού περιβλήματος, σφάλματα συγκολλήσεων).
- Καταστροφές από άγκυρες ή από δραστηριότητες ψαρέματος (συμβάντα τα οποία είτε προκαλούν άμεση καταστροφή είτε οδηγούν σε μηχανική αποτυχία ή αποτυχία γραμμής).

➪ *Προσέγγιση Andrews.*

Η έρευνα που πραγματοποιήθηκε από τον Andrews (1984) απέδειξε ότι η εκτίμηση της αξιοπιστίας ενός υποθαλάσσιου αγωγού χαρακτηρίζεται από την αβεβαιότητα των κινδύνων τους οποίους ενδέχεται να συναντήσει ο αγωγός. Μετά από μια σύντομη ανάλυση των αποτυχιών των παράκτιων αγωγών και

των αγωγών του Κόλπου του Μεξικού, ανέκυψαν δύο παράγοντες ως οι σημαντικότερες αιτίες των αποτυχιών. Οι παράγοντες αυτοί ήταν τρίτου βαθμού καταστροφή καθώς και η διάβρωση.

Η μελέτη επικεντρώθηκε στους αγωγούς της Βόρειας Θάλασσας και έδειξε ότι οι βλάβες εμφανίζονταν πιο συχνά κατά την φάση της κατασκευής και ήταν περισσότερο μηχανικής φύσης.

Ο Andrews, υπέδειξε επίσης ότι παρόλο που οι φυσικοί κίνδυνοι αποτελούν ένα σημαντικό πρόβλημα για τους υποθαλάσσιους αγωγούς, είναι εξαιρετικά δύσκολο να αναλυθούν και να ποσοτικοποιηθούν. Τέτοιου είδους κινδύνους αποτελούν οι αστάθεια τους εδάφους, η διάβρωση, οι σεισμοί και οι καταιγίδες.

Τελικά ο Andrews κατέληξε στο συμπέρασμα ότι η ευστάθεια ενός αγωγού θα έπρεπε να είναι μέρος κάθε εξίσωσης αξιοπιστίας, γεγονός όμως το οποίο αποτελεί έργο πολύ δύσκολο, και παραμένει ένα σημαντικό κενό στην εξίσωση αξιοπιστίας.

- *Προσέγγιση Cannon.*

Η μελέτη που δημοσιεύτηκε από τον Cannon (1985) περιέγραφε μια έρευνα αξιοπιστίας για συστήματα αγωγών πετρελαίου και φυσικού αερίου που λειτουργούν στον Βρετανικό τομέα της Βορείου Θάλασσας. Η ανάλυση των δεδομένων αφορούσε μόνο τις αποτυχίες των αγωγών οι οποίες και χωρίζονταν σε τρεις κύριες κατηγορίες:

1. Όλα τα αναφερόμενα γεγονότα. Στην κατηγορία αυτή περιλαμβάνονται όλα τα συμβάντα που απαιτούν λεπτομερή ανάλυση.
2. Γεγονότα που προκαλούν διακοπή λειτουργίας. Η κατηγορία αυτή περικλείει:
   - συμβάντα όπου η διακοπή λειτουργίας εμφανίστηκε είτε πριν την εμφάνιση της βλάβης είτε στο στάδιο επισκευής της βλάβης και
   - συμβάντα στα οποία συνέβη σοβαρή βλάβη.

   Τέλος, μερικά σημαντικά συμπεράσματα της μελέτης του Cannon είναι:

i. Το κύριο μέρος του ρυθμού αποτυχιών προερχόταν από βλάβες που οφείλονταν σε άγκυρες και δίκτυα,
ii. Τα δίκτυα αγωγών μήκους έως δύο μίλια είχαν μεγαλύτερο ρυθμό αποτυχιών και
iii. Οι αγωγοί με μικρότερες διαμέτρους παρουσίασαν μεγαλύτερο ρυθμό αποτυχιών.

➪ *Άλλες προσεγγίσεις και συμπεράσματα.*

Το Battelle Memorial Institute την περίοδο 1982-1983 δημοσίευσε μια μελέτη με τίτλο «Οι οικονομικές συνέπειες της αξιοπιστίας αγωγών». Αντικείμενο της μελέτης αυτής ήταν να προσδιορίσει τα χαρακτηριστικά γνωρίσματα των αγωγών τα οποία επηρεάζουν τις συχνότητες αποτυχιών τους και να δημιουργήσει, με βάση τα αρχεία των αγωγών και των αποτυχιών τους, μια σειρά συναρτήσεων οι οποίες εξασφαλίζουν εκτιμήσεις των συχνοτήτων αποτυχιών. Η ανάλυση αυτή περιορίστηκε σε συμβάντα που εμφανίστηκαν μετά την εγκατάσταση των αγωγών, αφού και σκοπός της μελέτης ήταν η βελτιστοποίηση του ελέγχου και της συντήρησης των εν λειτουργία αγωγών. Το πιο σημαντικό τμήμα του έργου αυτού υπήρξε η συλλογή κατάλληλων δεδομένων. Η συλλογή αυτή αποτέλεσε την πιο περιεκτική συγκέντρωση τέτοιου είδους πληροφοριών που είχε επιχειρηθεί μέχρι τότε.

Κεφάλαιο

2

ΔΙΑΧΩΡΙΣΤΙΚΗ ΑΝΑΛΥΣΗ

2.1 Γενικά

Με βάση τις βασικές ιδέες του διαχωρισμού, η διαχωριστική ανάλυση είναι μια στατιστική τεχνική η οποία επιτρέπει σε κάποιον που ενδιαφέρεται να μελετήσει τις διαφορές μεταξύ δύο ή περισσοτέρων ομάδων ατόμων, λαμβάνοντας υπόψη πολλά χαρακτηριστικά τους ταυτόχρονα. Επί παραδείγματι, δοθέντος μόνο του ύψους των ανθρώπων μιας περιοχής, κάποιος θα ήθελε να τους διαχωρίσει σε άνδρες και γυναίκες. Εντούτοις, μόνο το ύψος μπορεί να είναι ανεπαρκής τρόπος για τον διαχωρισμό του πληθυσμού μεταξύ των δύο φύλων. Όμως ένας συνδυασμός μετρήσεων του ύψους και της διαμέτρου της μέσης, θα μπορούσε να είναι επαρκής.

Η διαχωριστική συνάρτηση είναι αυτή που διευκολύνει την διαχωριστική μέθοδο. Συνήθως η διαχωριστική συνάρτηση είναι ένας γραμμικός συνδυασμός των βασικών χαρακτηριστικών όλων των ατόμων του πληθυσμού, αλλά υπάρχουν και άλλα συναρτησιακά σχήματα (π.χ. δευτεροβάθμιες, εκθετικές και λογαριθμικές συναρτήσεις).

Όταν το μέγεθος του πληθυσμού είναι μεγάλο, η μέτρηση των χαρακτηριστικών όλων των ατόμων του πληθυσμού καταλήγει σε μία χρονοβόρα και οικονομικά ασύμφορη διαδικασία. Για τον λόγο αυτό, χρησιμοποιείται η μέθοδος της δειγματοληψίας.

Στην συγκεκριμένη εργασία γίνεται χρήση της γραμμικής διαχωριστικής συνάρτησης. Δύο είναι κυρίως οι λόγοι που συνιστούν την επιλογή αυτή. Αφενός, η γραμμική διαχωριστική συνάρτηση βρίσκει απόλυτη εφαρμογή σε περιπτώσεις που εξετάζονται χαρακτηριστικά συνεχή ή ασυνεχή, γιατί είναι ανεξάρτητη από τις κατανομές των χαρακτηριστικών. Αφετέρου, η χρήση της γραμμικής διαχωριστικής συνάρτησης επιτρέπει την ευκολότερη ερμηνεία της συμμετοχής του κάθε χαρακτηριστικού στον διαχωρισμό, σε σχέση με κάθε άλλο συναρτησιακό σχήμα.

Η γραμμική διαχωριστική συνάρτηση, έχει την παρακάτω μορφή, όπως αυτή θα χρησιμοποιηθεί:

$$Z_m = U_0 + U_1X_{1m} + U_2X_{2m} + \ldots + U_iX_{im} \qquad (2.1)$$

όπου,

Z_m : Η τιμή της διαχωριστικής μεταβλητής για το άτομο m,

X_{im} : Η μέτρηση του i χαρακτηριστικού του ατόμου m,

U_i : Ο καλύτερος διαχωριστικός συντελεστής που επιδέχεται το χαρακτηριστικό i έτσι ώστε να προκύψει ο καλύτερος δυνατός διαχωρισμός.

Το αντικείμενο της διαχωριστικής ανάλυσης είναι να «ανακαλύψει» αυτά τα χαρακτηριστικά «X_i» και τους συντελεστές τους «U_i» οι οποίοι κάνουν τον τέλειο διαχωρισμό του πληθυσμού σε δύο διακεκριμένες ομάδες. Η μεθοδολογία που ακολουθείται αποτελείται από τα εξής στάδια:

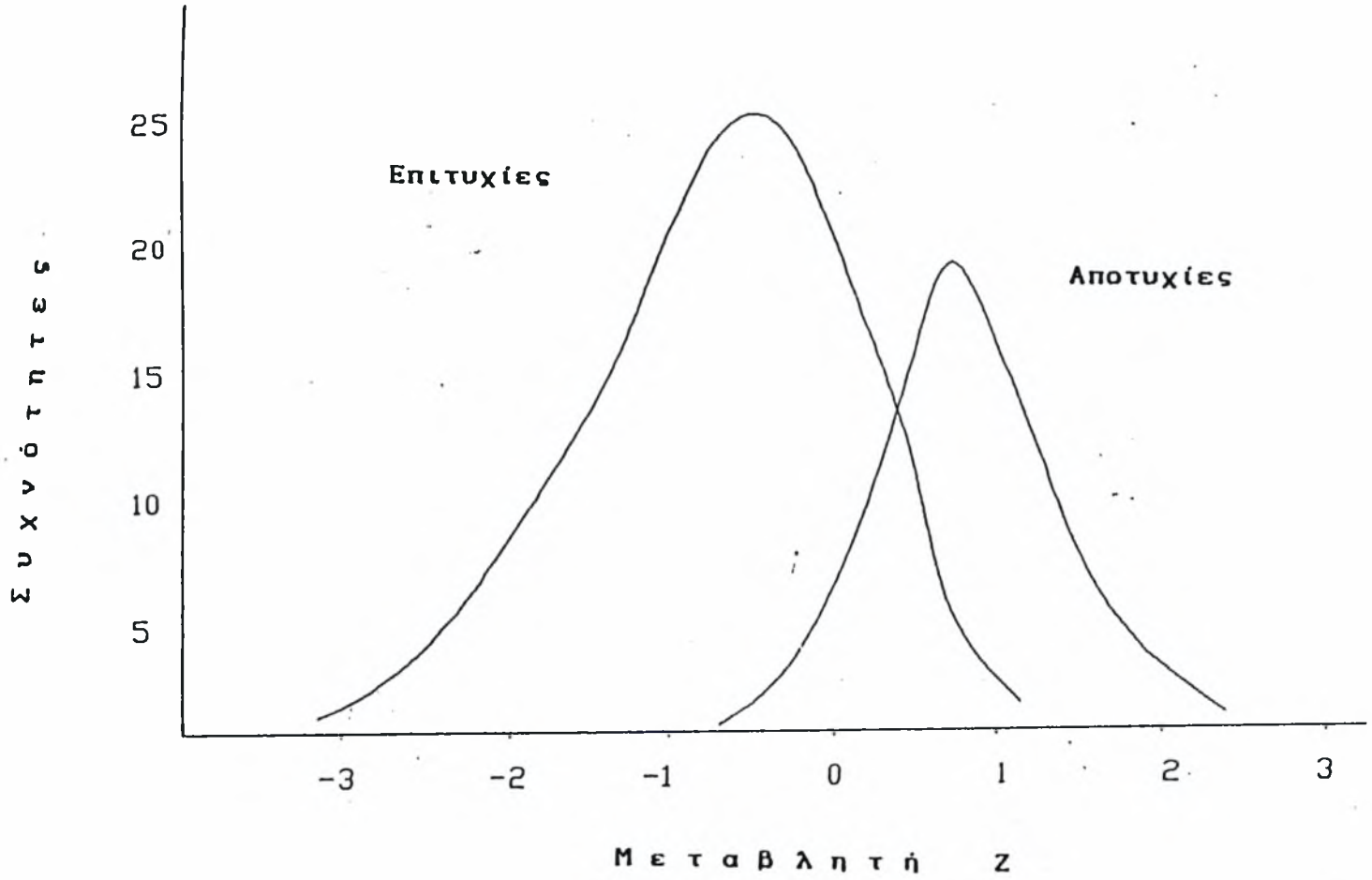
1. Παίρνονται αντιπροσωπευτικά δείγματα ατόμων από τις δύο διαχωριζόμενες ομάδες.
2. Παίρνονται οι τιμές X_i των χαρακτηριστικών i όλων των ατόμων.
3. Επιλέγονται εκείνα τα χαρακτηριστικά από τα οποία αναμένεται να προκύψει ο καλύτερος διαχωρισμός.

4. Υπολογίζονται οι τιμές των συντελεστών U_i που δίνουν τον καλύτερο διαχωρισμό με βάση την επιλογή που έγινε στο προηγούμενο στάδιο, και οι τιμές της μεταβλητής Z.
5. Επαναλαμβάνονται τα στάδια 3 και 4, χρησιμοποιώντας άλλους συνδυασμούς χαρακτηριστικών, ώσπου να επιτευχθεί ο καλύτερος, κατά το δυνατόν, διαχωρισμός των δειγμάτων.

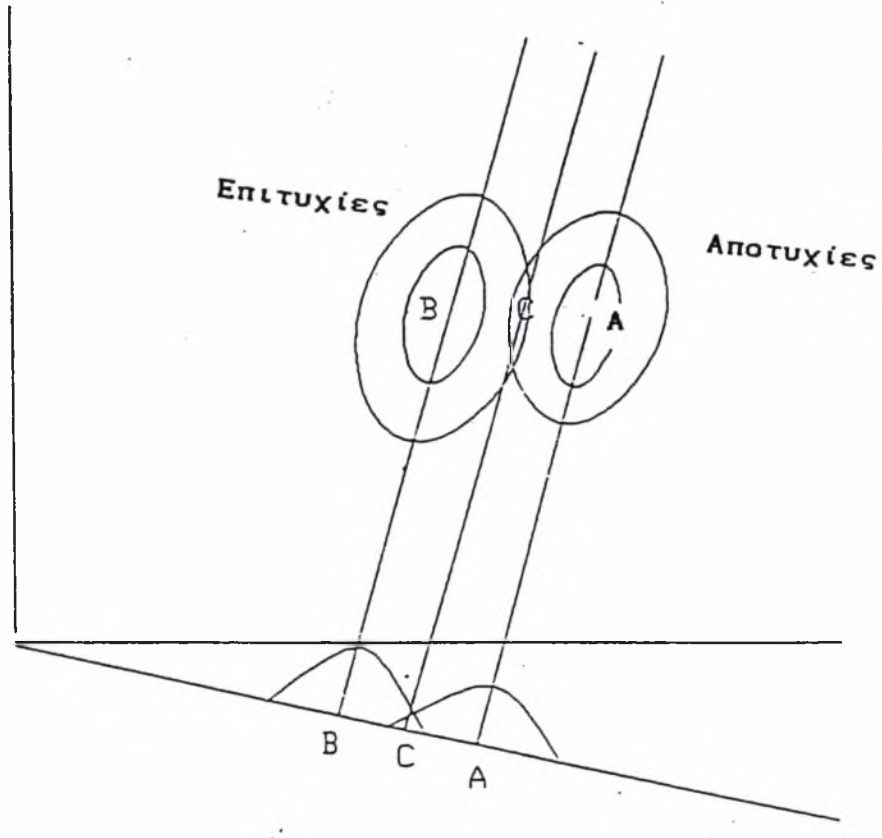
Επειδή κάθε άτομο έχει διαφορετικές τιμές X_i, για τα ίδια χαρακτηριστικά i, επακόλουθο είναι η τιμή της διαχωριστικής μεταβλητής να διαφέρει από άτομο σε άτομο, ακόμη και αν τα άτομα αυτά ανήκουν στο ίδιο δείγμα. Αποτέλεσμα αυτού είναι οι τιμές της Z να ακολουθούν διάφορα συναρτησιακά σχήματα τα οποία έχουν σχέση με τον τρόπο που είναι διασκορπισμένες οι τιμές των χαρακτηριστικών. Στις περισσότερες περιπτώσεις οι τιμές αυτές ακολουθούν την κανονική κατανομή.

Το αποτέλεσμα μιας τέτοιας ανάλυσης μπορεί να παρασταθεί γραφικά, στο σχήμα (2.1), όπου παρουσιάζονται οι τιμές της μεταβλητής Z και οι συχνότητές τους για δύο ομάδες. Προφανώς, ο σκοπός αυτής της ανάλυσης είναι να διαχωρίσει τις δύο κατανομές πυκνότητας πιθανότητας όσο το δυνατόν καλύτερα. Ο καλύτερος διαχωρισμός επιτυγχάνεται στην περίπτωση που οι κατανομές αυτές δεν επικαλύπτονται. Στην περίπτωση της επικάλυψης, όπως φαίνεται στο σχήμα, ο διαχωρισμός δεν είναι τέλειος γιατί τα άτομα που έχουν τιμή Z που βρίσκεται στην περιοχή επικάλυψης δεν διαχωρίζονται σωστά. Αυτό σημαίνει ότι αν γνωρίζαμε μόνο τις τιμές των χαρακτηριστικών δεν θα μπορούσαμε να καθορίσουμε από ποια ομάδα προέρχεται.

Στο σχήμα (2.2), δίνεται η γεωμετρική ερμηνεία της διαχωριστικής ανάλυσης στην οποία τα δύο σύνολα των ομοκεντρικών ελλείψεων αντιπροσωπεύουν, στην ιδεατή μορφή, τα άτομα των δύο ομάδων, στα οποία μετρούνται δύο χαρακτηριστικά. Κάθε έλλειψη είναι ο γεωμετρικός τόπος των σημείων ίσης πυκνότητας πιθανότητας (ή συχνότητας) μιας ομάδας. Για παράδειγμα, η εξωτερική έλλειψη της ομάδας των αποτυχιών, μπορεί να ορίζει την περιοχή στην οποία κείται το 90% των ατόμων και η εσωτερική έλλειψη, την περιοχή στην οποία βρίσκεται το 80%.



Σχήμα 2.1 Κατανομή πυκνότητας πιθανότητας της μεταβλητής Z



Σχήμα 2.2 Γεωμετρική ερμηνεία της διαχωριστικής ανάλυσης

2.2 Προϋποθέσεις

Προκειμένου να εφαρμοστεί η μαθηματική διαδικασία, όπως αυτή αναλύεται παρακάτω, πρέπει να ισχύουν ορισμένες βασικές προϋποθέσεις. Οι προϋποθέσεις αυτές είναι:

α. Απαιτούνται τουλάχιστον δύο ομάδες.
β. Κάθε ομάδα πρέπει να αποτελείται τουλάχιστον από δύο άτομα.
γ. Ένα χαρακτηριστικό που είναι γραμμικός συνδυασμός άλλων, πρέπει να παραλείπεται, γιατί δεν περιέχει καμία επιπρόσθετη πληροφορία.
δ. Δεν υπάρχει όριο στον αριθμό των χαρακτηριστικών, εκτός από το ότι ο αριθμός των ατόμων πρέπει να υπερβαίνει τον αριθμό των χαρακτηριστικών, τουλάχιστον κατά δύο.
ε. Οι τιμές των χαρακτηριστικών των ατόμων της κάθε ομάδας πρέπει να ακολουθούν την κανονική κατανομή ή διαφορετικά, κάθε ομάδα να ακολουθεί μια πολυμεταβλητή κανονική κατανομή, παρά το γεγονός ότι η γραμμική διαχωριστική συνάρτηση είναι ανεξάρτητη από την πολυκανονικότητα των ομάδων.
στ. Μια γραμμική διαχωριστική διαδικασία είναι ιδανική όταν οι μεταβλητότητες και οι συσχετίσεις των ανεξάρτητων χαρακτηριστικών σε μια ομάδα είναι ίδιες με τις αντίστοιχες των άλλων ομάδων.
ζ. Δεν πρέπει να υπάρχει αμφιβολία ως προς την ομάδα προέλευσης του κάθε ατόμου. Διαφορετικά η διαχωριστική συνάρτηση που θα προκύψει θα δώσει εσφαλμένα αποτελέσματα.

Όλες οι προϋποθέσεις που αναφέρονται ισχύουν φυσικά και στην περίπτωση της δειγματοληψίας, με την μόνη διαφορά ότι έχουμε δείγματα και όχι ομάδες. Ειδικότερα, για την προϋπόθεση (ε) θα πρέπει η κανονικότητα των χαρακτηριστικών να μπορεί να επεκταθεί και στις ομάδες.

2.3 Υπολογισμός της διαχωριστικής συνάρτησης

Μια εκτεταμένη περιγραφή της περίπλοκης και προχωρημένης μαθηματικής διαδικασίας του υπολογισμού της διαχωριστικής συνάρτησης και η θεωρητική της θεμελίωση, ξεφεύγει από τον σκοπό της παρούσας μελέτης. Η χρήση της ως υπολογιστικό εργαλείο, απαιτεί μόνο την κατανόηση των βασικών της αρχών και του τρόπου εφαρμογής της.

Στην συνέχεια, δίδονται όλοι οι απαραίτητες έννοιες και μαθηματικοί τύποι, καθώς και παρουσιάζεται μια εφαρμογή της μεθόδου.

Ορισμένοι συμβολισμοί οι οποίοι χρησιμοποιούνται παρακάτω, είναι:

g: ο αριθμός των ομάδων,

n_g: ο αριθμός των ατόμων της ομάδας g,

n: ο συνολικός αριθμός των ατόμων,

v: ο αριθμός των χαρακτηριστικών,

X_{igm}: η τιμή του χαρακτηριστικού i, του ατόμου m της ομάδας g,

M_{ig}: η μέση τιμή του χαρακτηριστικού i, των ατόμων της ομάδας g,

M_i: η μέση τιμή του χαρακτηριστικού i, σε ολόκληρο τον πληθυσμό.

Στην περίπτωση της δειγματοληψίας οι παραπάνω ορισμοί ισχύουν αν οι ομάδες αντικατασταθούν από τα δείγματα των ομάδων.

Μερικές στατιστικές τεχνικές είναι απαραίτητες προκειμένου να μετρηθεί ο βαθμός διασποράς των χαρακτηριστικών των ατόμων και οι συσχετίσεις μεταξύ τους. Ένας απλός πίνακας, με τις μέσες τιμές και τις τυπικές αποκλίσεις των χαρακτηριστικών των ατόμων των ομάδων, δεν είναι αρκετός, επειδή δεν θα έδειχνε τις συσχετίσεις των χαρακτηριστικών. Αντίθετα, ο πίνακας T[1], όπως παρουσιάζεται παρακάτω, μπορεί να χρησιμοποιηθεί σαν μια μέτρηση της διασποράς των χαρακτηριστικών από άτομο σε άτομο σε όλες τις ομάδες. Έτσι, κάθε στοιχείο T_{ij} του πίνακα T δίνεται από την ακόλουθη σχέση:

[1] Ο πίνακας T συναντάται στην αγγλική βιβλιογραφία σαν Total Sums of Squares and Cross Products matrix (SSCP).

$$T_{ij} = \sum_{g=1}^{g} \sum_{m=1}^{n_g} \left(X_{igm} - M_i \right) \left(X_{jgm} - M_j \right) \tag{2.2}$$

Οι όροι $(X_{igm} - M_i)$ και $(X_{jgm} - M_j)$ δείχνουν πόσο ένα χαρακτηριστικό αποκλίνει από την μέση τιμή. Αν i = j, τότε υπολογίζεται η διασπορά του χαρακτηριστικού και αν i≠j, τότε υπολογίζεται η συσχέτιση ανάμεσα σ' αυτά τα δύο χαρακτηριστικά.

Αν κάθε στοιχείο του πίνακα T διαιρεθεί με $(n - g)$, τότε προκύπτει ο πίνακας μεταβλητότητας - συμμεταβλητότητας για ολόκληρο τον πληθυσμό[1], που χρησιμοποιείται ευρέως στους στατιστικούς ελέγχους.

Τα στοιχεία ενός άλλου πίνακα, των συνολικών συντελεστών συσχέτισης, υπολογίζονται διαιρώντας κάθε στοιχείο του πίνακα T με την τετραγωνική ρίζα του γινομένου των δύο διαγωνίων στοιχείων, που βρίσκονται στην ίδια γραμμή και στήλη[2]. Μια υψηλή συσχέτιση ανάμεσα σε δύο χαρακτηριστικά σημαίνει ότι οι πληροφορίες που δίνονται χρησιμοποιώντας και τα δύο στην διαχωριστική συνάρτηση, δεν αλλάζουν αν χρησιμοποιήσουμε μόνο το ένα από αυτά.

Όταν οι μέσες τιμές των χαρακτηριστικών, των ατόμων των ομάδων, δεν ταυτίζονται, τότε οι διασπορές των χαρακτηριστικών σε κάθε ομάδα ξεχωριστά θα είναι μικρότερες από τις συνολικές διασπορές.

Αυτές μπορούν να υπολογισθούν από τον πίνακα W[3]. Κάθε στοιχείο W_{ij} του πίνακα W δίνεται από την εξίσωση:

$$W_{ij} = \sum_{g=1}^{g} \sum_{m=1}^{n_g} \left(X_{igm} - M_{ig} \right) \left(X_{jgm} - M_{jg} \right) \tag{2.3}$$

[1] Ο πίνακας αυτός συναντάται στην αγγλική βιβλιογραφία σαν Total Variance - Covariance Matrix ή σαν Dispersion Matrix.

[2] Συναντάται στην διεθνή βιβλιογραφία σαν Total Correlation - Coefficient Matrix.

[3] Συναντάται στην διεθνή βιβλιογραφία σαν Within - Groups Sums of Squares and Cross Products Matrix (WSSCP).

Τώρα, είναι εύκολο να μετατραπεί ο πίνακας W σε έναν πίνακα μεταβλητότητας - συμμεταβλητότητας ή σε έναν πίνακα συντελεστών συσχέτισης με την ίδια διαδικασία που περιγράφηκε για τον πίνακα T.

Όταν οι μέσες τιμές των ομάδων είναι ίσες, τότε οι πίνακες T και W είναι ίδιοι. Αν οι μέσες τιμές είναι διαφορετικές, υπάρχει μια διαφορά ανάμεσα στον T και τον W και εκφράζεται από τον πίνακα B[1], όπου $B = T - W$. Τώρα οι πίνακες W και B περιέχουν όλες τις βασικές πληροφορίες για τις ομάδες ξεχωριστά και για τις σχέσεις μεταξύ τους. Ο Fisher πρότεινε ένα κριτήριο για την μέτρηση της διαφοράς των ομάδων. Η μαθηματική διατύπωση αυτού του κριτηρίου επεκτάθηκε από άλλους ερευνητές και προέκυψε η εξίσωση:

$$\left(W^{-1}B - \lambda I\right)V = 0 \qquad (2.4)$$

όπου:

λ: μιας από τις ιδιοτιμές του πίνακα $W^{-1}B$,

I: ο μοναδιαίος πίνακας,

V: το ιδιοδιάνυσμα που προκύπτει από την ιδιοτιμή λ,

0: ο μηδενικός πίνακας.

Η λύση της εξίσωσης (2.4), περιλαμβάνει τις ιδιοτιμές λ του πίνακα $W^{-1}B$ και τα ιδιοδιανύσματα που αντιστοιχούν σ' αυτές τις ιδιοτιμές, και επιλύει το πρόβλημα τις μεγιστοποίησης της διαφοράς των ομάδων. Οι ιδιοτιμές αυτές είναι οι λύσεις της επόμενης εξίσωσης:

$$det\left(W^{-1}B - \lambda I\right) = 0 \qquad (2.5)$$

Από τις ρίζες τις εξίσωσης απορρίπτονται οι μιγαδικές και οι αρνητικές, επειδή σ' αυτές δεν μπορεί να δοθεί καμία στατιστική σημασία. Για κάθε θετική ιδιοτιμή προκύπτει ένα σύστημα της παρακάτω μορφής, από το οποίο υπολογίζεται το ιδιοδιάνυσμα (V_1, V_2, V_3, ..., V_v).

[1] Συναντάται στην διεθνή βιβλιογραφία σαν Between - Groups Sums of Squares and Cross Products Matrix (BSSCP).

$$
\begin{aligned}
&(B_{11} - \lambda W_{11})V_1 + (B_{12} - \lambda W_{12})V_2 + \ldots + (B_{1v} - \lambda W_{1v})V_v = 0 \\
&(B_{21} - \lambda W_{21})V_1 + (B_{22} - \lambda W_{22})V_2 + \ldots + (B_{2v} - \lambda W_{2v})V_v = 0 \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&(B_{v1} - \lambda W_{v1})V_1 + (B_{v2} - \lambda W_{v2})V_2 + \ldots + (B_{vv} - \lambda W_{vv})V_v = 0
\end{aligned}
\tag{2.6}
$$

Το παραπάνω σύστημα όπως εύκολα αποδεικνύεται είναι αόριστο. Μία συνήθης πρακτική που ακολουθείται για την επίλυση ενός τέτοιου συστήματος είναι να δίνουμε σε κάποιο άγνωστο την τιμή της μονάδας.

Επειδή οι συντελεστές (V_1, V_2, V_3, ..., V_v) είναι τελείως ανερμήνευτοι και οι τιμές της μεταβλητής Z που θα προέκυπταν δεν θα είχαν καμία σημασία, χρησιμοποιούνται στην εξίσωση (2.1) οι συντελεστές (U_0, U_1, U_2, ..., U_v), που υπολογίζονται σύμφωνα με τις ακόλουθες σχέσεις:

$$
\begin{aligned}
U_i &= \frac{V_i\sqrt{n-g}}{\sqrt{VTV^t}} \\
U_0 &= -\sum_{i=1}^{v} U_i M_i
\end{aligned}
\tag{2.7}
$$

Οι συντελεστές U_i δείχνουν την απόλυτη συνεισφορά κάθε χαρακτηριστικού στον υπολογισμό της διαχωριστικής μεταβλητής Z σύμφωνα με την εξίσωση (2.1). Με την χρησιμοποίηση των U_i, οι τιμές της διαχωριστικής μεταβλητής που υπολογίζονται για όλα τα άτομα έχουν μέση τιμή μηδέν και τυπική απόκλιση μονάδα.

Όπως γίνεται αντιληπτό, προκύπτουν πολλές διαχωριστικές συναρτήσεις και επομένως τίθεται το ζήτημα της επιλογής εκείνης της συνάρτησης που έχει την μεγαλύτερη διαχωριστική ισχύ. Αυτό σημαίνει ότι η περιοχή επικάλυψης των κατανομών της μεταβλητής Z είναι η μικρότερη δυνατή, άρα οι κατανομές είναι κατά το μέγιστο δυνατό απομακρυσμένες (Σχ. 2.1).

2.4 Ερμηνεία των διαχωριστικών συντελεστών

Οι διαχωριστικοί συντελεστές U_i, που υπολογίστηκαν προηγουμένως μπορούν να οδηγήσουν σε ορισμένα συμπεράσματα σχετικά με την διαχωριστική ισχύ που φέρει κάθε χαρακτηριστικό. Έτσι, είναι δυνατόν να προκύψουν χαρακτηριστικά χωρίς διαχωριστική ισχύ τα οποία εξαιρούνται από τα επόμενα στάδια της ανάλυσης. Επίσης, από αυτούς τους συντελεστές μπορεί να προσδιοριστεί η βαρύτητα κάθε χαρακτηριστικού στον διαχωρισμό των ομάδων.

Υπάρχουν διάφορα κριτήρια με τα οποία αναγνωρίζουμε αν κάποια χαρακτηριστικό έχει διαχωριστική ισχύ. Στην περίπτωση του διαχωρισμού δύο ομάδων, ένα κριτήριο που μπορούμε να χρησιμοποιήσουμε είναι:

$$U_i\left(M_{i1} - M_{i2}\right)\left(Z_{m1} - Z_{m2}\right) > 0 \tag{2.8}$$

όπου:

Z_{m1}: η μέση τιμή της μεταβλητής Z της ομάδας 1

Z_{m2}: η μέση τιμή της μεταβλητής Z της ομάδας 2

Για να προκύψει το κριτήριο (2.8) γίνονται οι ακόλουθες παραδοχές:

- Ένα χαρακτηριστικό έχει διαχωριστική ισχύ όταν συμβάλλει στην απομάκρυνση των κατανομών της μεταβλητής Z των δύο ομάδων. Αυτό εκφράζεται με την αύξηση της διαφοράς των μέσων τιμών Z_{m1} και Z_{m2} των κατανομών.
- Οι μέσες τιμές υπολογίζονται από τις επόμενες εξισώσεις:

$$Z_{mg} = U_0 + U_1 M_{1g} + \ldots + U_v M_{vg} \qquad g = 1,2$$

άρα η διαφορά των μέσων τιμών των κατανομών ισούται με

$$\left(Z_{m1} - Z_{m2}\right) = U_1\left(M_{11} - M_{12}\right) + \ldots + U_v\left(M_{v1} - M_{v2}\right)$$

- Ένα χαρακτηριστικό i έχει διαχωριστική ισχύ όταν το πρόσημο του όρου $U_i(M_{i1}-M_{i2})$ είναι το ίδιο με το πρόσημο της διαφοράς $(Z_{m1}-Z_{m2})$. Κάθε τέτοιο χαρακτηριστικό ονομάζεται ευσταθές. Στην αντίθετη περίπτωση το χαρακτηριστικό δεν έχει καθόλου διαχωριστική ισχύ και ονομάζεται ασταθές.

Ανεξάρτητα από το αν ένα χαρακτηριστικό είναι ευσταθές ή ασταθές, επιδρά στον διαχωρισμό με κάποια βαρύτητα. Η βαρύτητα αυτή δεν μπορεί να προκύψει απ' ευθείας από τους συντελεστές U_i επειδή είναι εξαρτημένοι από την κλίμακα των μετρήσεων των παραπάνω χαρακτηριστικών. Ο μόνος τρόπος για να ανεξαρτητοποιηθούν οι συντελεστές αυτοί από τις μονάδες μέτρησης είναι η κανονικοποίησή τους. Οι κανονικοποιημένοι συντελεστές μπορούν να υπολογιστούν κάνοντας χρήση της ακόλουθης σχέσης:

$$S_i = U_i \sqrt{\frac{W_{ii}}{n-g}} \tag{2.9}$$

όπου W_{ii} είναι το διαγώνιο στοιχείο του πίνακα W που αντιστοιχεί στο χαρακτηριστικό i και $W_{ii}/(n-g)$ είναι η μεταβλητότητα του σε ολόκληρο τον πληθυσμό. Όσο μεγαλύτερος, κατ' απόλυτη τιμή, είναι ένας κανονικοποιημένος συντελεστής τόσο μεγαλύτερη είναι και η βαρύτητά του.

2.5 Έλεγχος της διαχωριστικής συνάρτησης

Η επιτυχία κάθε διαχωριστικής διαδικασίας μπορεί να ελεγχθεί χρησιμοποιώντας δύο διαφορετικές στατιστικές τεχνικές, όπως περιγράφονται πιο κάτω. Με την πρώτη στατιστική τεχνική ελέγχεται η ποιότητα των δεδομένων που διατίθενται από τις ομάδες ή από τα δείγματα, στην περίπτωση της δειγματοληψίας. Επειδή από τις ομάδες ή τα δείγματα προκύπτουν πολλές διαχωριστικές συναρτήσεις, χρησιμοποιείται η δεύτερη στατιστική τεχνική προκειμένου να επιλεγεί η καλύτερη από αυτές. Όπως γίνεται αντιληπτό αν τα αποτελέσματα

της πρώτης τεχνικής δεν κριθούν ως ικανοποιητικά η εφαρμογή της δεύτερης είναι άνευ σημασίας.

Α. Η ***πρώτη τεχνική*** μετράει τον διαχωρισμό των ομάδων, που επιτυγχάνεται από τα χαρακτηριστικά που έχουν επιλεχθεί, σαν ένα λόγο οριζουσών. Ο λόγος αυτός ονομάζεται λόγος οριζουσών του Wilk, συμβολίζεται με «Λ» και ορίζεται ως:

$$\Lambda = \frac{|W|}{|T|}$$

όπου ο αριθμητής είναι η ορίζουσα του πίνακα W και ο παρανομαστής η ορίζουσα του πίνακα T. Αν η τιμή του Λ προσεγγίζει το μηδέν, τότε οι τιμές των ομάδων είναι χωρισμένες κατά το μέγιστο και ο διαχωρισμός είναι πολύ καλός. Στην περίπτωση που ο υπολογισμός της διαχωριστικής συνάρτησης έχει γίνει με την χρήση δειγμάτων που αντιπροσωπεύουν τις ομάδες, τότε η χρήση του Wilk's Λ είναι περισσότερο ένα βήμα για τον έλεγχο της στατιστικής σημαντικότητας της απόστασης ανάμεσα στις μέσες τιμές των ομάδων, παρά ένα τελικό αποτέλεσμα που θα μπορούσε να δώσει μια μέτρηση του διαχωρισμού. Αυτό σημαίνει ότι δεν είμαστε σε θέση να γνωρίζουμε αν η τιμή του Wilk's Λ που βρέθηκε από τα συγκεκριμένα δείγματα μπορεί να επεκταθεί σε ολόκληρο τον πληθυσμό πριν κάνουμε τον έλεγχο υποθέσεων που περιγράφεται παρακάτω.

Αρχικά, διατυπώνουμε τις ακόλουθες υποθέσεις:

- Μηδενική υπόθεση (H_0): οι μέσες τιμές της μεταβλητής Z των δύο ομάδων είναι ίσες, δηλαδή Wilk's Λ = 1 .
- Εναλλακτική υπόθεση (H_1): οι μέσες τιμές της μεταβλητής Z των δύο ομάδων δεν είναι ίσες.

Για να κάνουμε τον έλεγχο υποθέσεων είναι απαραίτητο να γνωρίζουμε την κατανομή που ακολουθεί το Wilk's Λ. Έχει βρεθεί ότι το Wilk's Λ ανήκει σε μια οικογένεια καμπυλών με τρεις παραμέτρους που είναι συναρτήσεις του αριθμού των ομάδων, του αριθμού των ατόμων και του αριθμού των χαρακτηριστικών.

Πρόσφατα, δημιουργήθηκαν πίνακες που περιέχουν το Wilk's Λ για ορισμένες ειδικές τιμές των παραμέτρων του. Στην γενική της χρήση όμως, η κατανομή Wilk's Λ μετασχηματίζεται προσεγγιστικά σε κατανομή X^2 ή σε κατανομή F. Οι μελετητές που ασχολήθηκαν μ' αυτούς τους μετασχηματισμούς κατέληξαν στο συμπέρασμα ότι η προσέγγιση που γίνεται με την κατανομή F είναι καλύτερη από αυτήν της κατανομής X^2, επειδή δίνει αποτελέσματα που είναι συμβατά με την κατανομή του Wilk's Λ ακόμη και για πολύ μικρό αριθμό βαθμών ελευθερίας. Στην συγκεκριμένη εργασία, γίνεται χρήση της κατανομής F, με βάση την ανάπτυξη αυτής από τον Rao. Στην προσέγγιση αυτή είναι απαραίτητο να υπολογιστούν οι τιμές των παρακάτω εξισώσεων:

$$F = \frac{1-\Lambda^{\frac{1}{s}}}{\Lambda^{\frac{1}{s}}} \cdot \frac{ms - \frac{v(g-1)}{2} + 1}{v(g-1)} \tag{2.10}$$

όπου:

$$m = n - 1 - (v+g)/2$$

$$s = \sqrt{\frac{v^2(g-1)^2 - 4}{v^2 + (g-1)^2 - 5}}$$

με βαθμούς ελευθερίας:

$$v_1 = v(g-1) \qquad\qquad v_2 = ms - v(g-1)/2 + 1$$

και όπου Λ συμβολίζεται το Wilk's Λ.

Με τους βαθμούς ελευθερίας v_1 και v_2 μπορούμε, με την βοήθεια των πινάκων της κατανομής F, για ένα ορισμένο επίπεδο σημαντικότητας α%, να διαβάσουμε μια τιμή της μεταβλητής F.

Εάν η τιμή είναι μικρότερη από αυτήν που υπολογίσαμε από την σχέση 2.10, τότε απορρίπτουμε την μηδενική υπόθεση H_0 και δεχόμαστε ότι η τιμή του Wilk's Λ που βρήκαμε ισχύει για όλο τον πληθυσμό. Η πιθανότητα λαθεμέ-

νης απόφασης είναι α%, γνωστό στην θεωρία των πιθανοτήτων και ως σφάλμα α' είδους. Όπως γίνεται σαφές, όσο μικρότερο είναι το επίπεδο σημαντικότητας τόσο μικρότερη είναι η πιθανότητα να λάβουμε λανθασμένη απόφαση.

Στην περίπτωση που η τιμή που διαβάζουμε από τον πίνακα είναι μεγαλύτερη από αυτήν που υπολογίσαμε τότε δεχόμαστε την μηδενική υπόθεση H_0. Όμως, αυτό που ενδιαφέρει είναι να ισχύει η τιμή του Wilk's Λ για όλο τον πληθυσμό. Στο σημείο αυτό θα πρέπει να ερευνηθεί να η αποδοχή της H_0 εξαρτάται από τυχαίους παράγοντες που οφείλονται σε διακυμάνσεις της δειγματοληψίας. Εάν μπορούμε να αποδώσουμε το γεγονός αυτό στις διακυμάνσεις της δειγματοληψίας, τότε η τιμή του Wilk's Λ = 1 είναι πραγματική και οποιοσδήποτε χωρισμός ανάμεσα στις ομάδες είναι αδύνατος, με βάση τα χαρακτηριστικά που έχουν εξετασθεί. Ο μόνος τρόπος για να συνεχίσουμε την ανάλυση είναι να επαναλάβουμε την διαδικασία με διαφορετικό συνδυασμό χαρακτηριστικών.

Ο παραπάνω έλεγχος της μηδενικής υπόθεσης της ισότητας των μέσων τιμών, προϋποθέτει ότι οι πίνακες μεταβλητότητας - συμμεταβλητότητας των g ομάδων, υπολογίζονται από δείγματα που προέρχονται από g πολυμεταβλητές κανονικές ομάδες με την ίδια διασπορά. Ουσιαστικά, αυτή είναι μια βασική προϋπόθεση για την εφαρμογή της γραμμικής διαχωριστικής συνάρτησης. Ο Box ανέπτυξε ένα κριτήριο (Box's B) για τον έλεγχο της παραπάνω προϋπόθεσης. Άρα, πρέπει να εφαρμόσουμε και πάλι την θεωρία του ελέγχου υποθέσεων, με βάση τις παρακάτω υποθέσεις:

- Μηδενική υπόθεση (H_0): οι πίνακες μεταβλητότητας - συμμεταβλητότητας των ομάδων είναι ίσοι.
- Εναλλακτική υπόθεση (H_1): οι πίνακες μεταβλητότητας - συμμεταβλητότητας των ομάδων δεν είναι ίσοι.

Σύμφωνα με το κριτήριο αυτό, αρχικά υπολογίζεται η μεταβλητή:

$$B = (n-g)\,ln|W| - \sum_{g=1}^{g}(n_g - 1)\,ln|W_g| \qquad (2.11)$$

όπου:

|W| : η ορίζουσα του πίνακα μεταβλητότητας - συμμεταβλητότητας του συνολικού δείγματος.

$|W_g|$: η ορίζουσα του πίνακα μεταβλητότητας - συμμεταβλητότητας του δείγματος που προέρχεται από την ομάδα g.

Κάθε στοιχείο του πίνακα W_g υπολογίζεται από την ακόλουθη σχέση:

$$\left(W_g\right)_{ij} = \frac{1}{n_g - 1}\sum_{m=1}^{n_g}\left(X_{igm} - M_{ig}\right)\left(X_{jgm} - M_{jg}\right) \qquad (2.12)$$

Αν i = j τότε τα στοιχεία του πίνακα είναι οι μεταβλητότητες των χαρακτηριστικών i και αν i ≠ j τότε τα στοιχεία του πίνακα είναι οι συμμεταβλητότητες των χαρακτηριστικών i , j.

Παρατηρώντας την προηγούμενη εξίσωση, συμπεραίνουμε ότι για να ισχύει η υπόθεση H_0 θα πρέπει η μεταβλητή B να είναι όσο γίνεται μικρότερη. Επειδή δεν είμαστε σε θέση να γνωρίζουμε την κατανομή που ακολουθεί η μεταβλητή B, την μετασχηματίζουμε σε μια μεταβλητή (Box's F) που ακολουθεί την κατανομή F. Ο μετασχηματισμός γίνεται σύμφωνα με την παρακάτω διαδικασία:

- Υπολογίζονται οι παράμετροι:

$$A_1 = \left(\sum_{g=1}^{g}\frac{1}{n_g - 1} - \frac{1}{(n-g)}\right)\cdot\frac{2v^2 + 3v - 1}{6(g-1)(v+1)}$$

$$A_2 = \left(\sum_{g=1}^{g}\frac{1}{\left(n_g - 1\right)^2} - \frac{1}{(n-g)^2}\right)\cdot\frac{(v-1)(v+2)}{6(g-1)}$$

- Υπολογίζονται οι βαθμοί ελευθερίας και η τιμή της μεταβλητής F κατά περίπτωση.

i) Αν $A_2 - A_1^2 > 0$ τότε

$$\nu_1 = \frac{(g-1)\nu(\nu+1)}{2}$$

$$\nu_2 = \frac{\nu_1 + 2}{A_2 - A_1^2}$$

$$b = \frac{\nu_1}{1 - A_1 - \frac{\nu_1}{\nu_2}} \qquad (2.13)$$

$$F = \frac{B}{b}$$

ii) Αν $A_2 - A_1^2 < 0$ τότε

$$\nu_1 = \frac{(g-1)\nu(\nu+1)}{2}$$

$$\nu_2 = \frac{\nu_1 + 2}{-A_2 + A_1^2}$$

$$b = \frac{\nu_2}{1 - A_1 - \frac{2}{\nu_2}} \qquad (2.14)$$

$$F = \frac{\nu_2 B}{\nu_1 (b - B)}$$

Με παρόμοιο τρόπο, όπως αναφέρθηκε στον προηγούμενο έλεγχο υποθέσεων, μπορούμε να απορρίψουμε ή να αποδεχτούμε την μηδενική υπόθεση H_0. Αντίθετα, όμως, με τον προηγούμενο έλεγχο, εδώ επιδιώκουμε την αποδοχή της μηδενικής υπόθεσης H_0. Βέβαια και στην περίπτωση αυτή η πιθανότητα λανθασμένης απόφασης είναι α% (σφάλμα α' είδους).

Αν απορρίψουμε την μηδενική υπόθεση H_0, τότε πρέπει να διαπιστώσουμε κατά πόσο η απόρριψη αυτή οφείλεται σε τυχαίους παράγοντες ή πρόκειται για ένα πραγματικό γεγονός. Αν όντως οι πίνακες μεταβλητότητας - συμμεταβλητότητας δεν είναι ίσοι, τότε η γραμμική διαχωριστική συνάρτηση δεν κάνει τον καλύτερο δυνατό διαχωρισμό, αφού μια προϋπόθεση για την εφαρμογή της δεν ισχύει. Μια λογική σκέψη είναι να εγκαταλειφθεί αυτό το συναρτησιακό σχήμα και να υιοθετηθεί κάποιο άλλο, όπως για παράδειγμα η τετραγωνική διαχωριστική συνάρτηση. Παρά το γεγονός ότι η χρήση της συνάρτησης αυτής

δεν απαιτεί την ισότητα των πινάκων που έχουν αναφερθεί, εντούτοις μικρές αποκλίσεις από την πολυκανονικότητα των ομάδων επηρεάζουν ιδιαίτερα την εφαρμοσιμότητά της. Έτσι, υπάρχουν περιπτώσεις που γίνεται χρήση της γραμμικής διαχωριστικής συνάρτησης, χωρίς να ισχύει η προϋπόθεση που αναφέρθηκε.

Β. Η ***δεύτερη στατιστική τεχνική*** μετράει την σχέση ανάμεσα στις ομάδες και στις διαχωριστικές συναρτήσεις, δηλαδή την ποσότητα του διαχωρισμού που φέρει η κάθε συνάρτηση. Επειδή κάθε διαχωριστική συνάρτηση προκύπτει από μια συγκεκριμένη ιδιοτιμή, το πρόβλημα ανάγεται τελικά στον προσδιορισμό της καλύτερης ιδιοτιμής. Η επιλογή αυτής γίνεται με βάση τον κανονικό συντελεστή συσχέτισης (Canonical Correlation Coefficient - C.C.C.) που δίνεται από την σχέση:

$$R_i = \sqrt{\frac{\lambda_i}{1+\lambda_i}} \tag{2.15}$$

Ο κανονικός συντελεστής συσχέτισης προέρχεται από μια στατιστική τεχνική η οποία ονομάζεται Ανάλυση Κανονικής Συσχέτισης (Canonical Correlation Analysis). Αν ο συντελεστής συσχέτισης προσεγγίζει την μονάδα, η σχέση ανάμεσα στην διαχωριστική συνάρτηση και στις ομάδες είναι πολύ υψηλή.

2.6 Επιλογή των χαρακτηριστικών της επόμενης διαχωριστικής συνάρτησης

Η επιλογή του καλύτερου συνδυασμού των ανεξάρτητων χαρακτηριστικών, από τον οποίο θα προκύψει η τελική διαχωριστική συνάρτηση, μπορεί να γίνει με δύο μεθόδους, την «ευθεία» και την «βηματική» (εμπειρική)[1] μέθοδο.

Η *«ευθεία»* μέθοδος είναι εκείνη κατά την οποία όλα τα ανεξάρτητα χαρακτηριστικά εισάγονται στην ανάλυση. Έτσι, η διαχωριστική συνάρτηση προκύπτει από όλα τα ανεξάρτητα χαρακτηριστικά, χωρίς να λαμβάνεται υπόψη η

διαχωριστική ισχύς του καθενός. Η «ευθεία» μέθοδος ενδείκνυται όταν είναι επιθυμητό να υπάρχουν όλα τα χαρακτηριστικά στην ανάλυση και δεν ενδιαφέρει η χρησιμοποίηση των αποτελεσμάτων που βασίζονται σε υποσυνδυασμούς των χαρακτηριστικών.

Η *«βηματική»* μέθοδος είναι εκείνη όπου τα ανεξάρτητα χαρακτηριστικά που χρησιμοποιούνται στην ανάλυση επιλέγονται με βάση την διαχωριστική τους ισχύ. Μερικές φορές ολόκληρος ο συνδυασμός μπορεί να δυσχεραίνει τον διαχωρισμό των ομάδων ή μερικά από τα χαρακτηριστικά να μην είναι χρήσιμα για τον διαχωρισμό των ομάδων. Επιλέγοντας τους επόμενους καλύτερους διαχωριστές (χαρακτηριστικά), προκύπτει ένας συνδυασμός χαρακτηριστικών ο οποίος είναι σχεδόν τόσο καλός ή και καλύτερος από ολόκληρο τον συνδυασμό, εξαιτίας της πιθανής ύπαρξης ασταθών χαρακτηριστικών. Κατά την εισαγωγή καινούργιων χαρακτηριστικών μέσα στον συνδυασμό, μερικά χαρακτηριστικά που υπάρχουν στο προηγούμενο βήμα μπορεί να χάσουν την διαχωριστική τους ισχύ. Αυτό συμβαίνει επειδή οι πληροφορίες που δίνουν για τις διαφορές των ομάδων είναι τώρα διαθέσιμες από κάποια άλλα χαρακτηριστικά. Η χρήση τέτοιων χαρακτηριστικών στην ανάλυση θεωρείται πλεονασμός και για τον λόγο αυτόν πρέπει να αποφεύγεται.

Υπάρχουν διάφορα κριτήρια επιλογής των χαρακτηριστικών, που βοηθούν στην εφαρμογή της «βηματικής» μεθόδου. Το κριτήριο που χρησιμοποιείται κυρίως, είναι το Wilk's Lambda (Λ). Όπως έχει προαναφερθεί, με το Wilk's Λ γίνεται ο ολικός πολυμεταβλητός έλεγχος F, για την διαφορά των μέσων τιμών της μεταβλητής Z των δύο ομάδων. Τα χαρακτηριστικά που μεγιστοποιούν τον λόγο F, ελαχιστοποιούν επίσης το Wilk's Λ, το οποίο αποτελεί μια μέτρηση του διαχωρισμού των ομάδων. Αυτό σημαίνει ότι όσο μικρότερη είναι η τιμή του Wilk's Λ, τόσο καλύτερος είναι ο διαχωρισμός που επιτυγχάνεται.

[1] Ο όρος «βηματική» μέθοδος είναι μια ελεύθερη μετάφραση της stepwise method που συναντάται στην αγγλική βιβλιογραφία.

Κεφάλαιο

3

ΤΑΞΙΝΟΜΗΣΗ

3.1 Γενικά

Στο προηγούμενο κεφάλαιο εξετάστηκε πως τα υπάρχοντα δεδομένα μπορούν να χωριστούν καλύτερα χρησιμοποιώντας μια γραμμική διαχωριστική συνάρτηση. Αν και η ανάλυση αυτή είναι ιδιαίτερα χρήσιμη για την κατανόηση των αιτίων που οδηγούν τα άτομα στο να ανήκουν σε κάποιες συγκεκριμένες ομάδες, και των χαρακτηριστικών που επιδρούν περισσότερο στον διαχωρισμό, ακόμη μεγαλύτερης πρακτικής σημασίας είναι η μελέτη των κανόνων ταξινόμησης, που μπορούν να χρησιμοποιηθούν για την τοποθέτηση ενός νέου ατόμου στην σωστή ομάδα. Τότε, θα είναι δυνατόν να προβλεφθεί αν ένα συγκεκριμένο άτομο έχει, ή όχι, υψηλή πιθανότητα να ανήκει σε μία από τις ομάδες, στο τέλος ενός προκαθορισμένου χρονικού ορίζοντα.

Το θέμα αυτό είναι στενά συνδεδεμένο με την προηγούμενη διαχωριστική ανάλυση, αφού η ικανότητα τοποθέτησης εξαρτάται από τα επιλεγμένα διαχωριστικά χαρακτηριστικά. Επιπλέον, εφαρμόζοντας τους κανόνες ταξινόμησης για τα άτομα των δειγμάτων από τα οποία προέκυψε η διαχωριστική συνάρτη-

ση, γίνεται ένας έλεγχος του διαχωρισμού των ομάδων και των δυνατοτήτων της ταξινόμησης για την κατάταξη νέων ατόμων.

Οι κανόνες ταξινόμησης που θα αναπτυχθούν, στοχεύουν στην κατάταξη των νέων ατόμων σε δύο ομάδες, αλλά μπορούν να επεκταθούν και για περισσότερες. Πρέπει να σημειωθεί ότι σπάνια παράγεται μια μέθοδος ταξινόμησης αλάνθαστη, επειδή δεν μπορεί να υπάρξει απόλυτη διαφορά ανάμεσα στα μετρούμενα χαρακτηριστικά των ατόμων. Αυτό συμβαίνει επειδή οι ομάδες συχνά επικαλύπτονται. Έτσι, είναι πιθανό να προβλεφθεί ότι ένα άτομο θα ανήκει σε μια ομάδα Α, ενώ στο τέλος του χρονικού ορίζοντα το άτομο βρέθηκε στην άλλη ομάδα Β.

3.2 Κανόνες ταξινόμησης

Η διαδικασία κατάταξης των ατόμων στις ομάδες βασίζεται σε κανόνες που σκοπό έχουν να ελαχιστοποιήσουν το συνολικό κόστος λανθασμένης ταξινόμησης ή κάτω από ορισμένες προϋποθέσεις, την ίδια την πιθανότητα λανθασμένης ταξινόμησης. Παρακάτω δίνονται ορισμένοι βασικοί συμβολισμοί:

X : Το διάνυσμα (vx1) των μετρούμενων διαχωριστικών χαρακτηριστικών (X_1, X_2, ..., X_v).

$f_1(X)$: Η συνάρτηση πυκνότητας πιθανότητας της (vx1) τυχαίας μεταβλητής X για την ομάδα 1.

$f_2(X)$: Η συνάρτηση πυκνότητας πιθανότητας της (vx1) τυχαίας μεταβλητής X για την ομάδα 2.

c(1/2) : Το κόστος λαθεμένης ταξινόμησης ενός ατόμου στην ομάδα 1, ενώ ανήκει στην ομάδα 2.

c(2/1) : Το κόστος λαθεμένης ταξινόμησης ενός ατόμου στην ομάδα 2, ενώ ανήκει στην ομάδα 1.

P(1/2) : Η πιθανότητα της λαθεμένης ταξινόμησης ενός ατόμου στην ομάδα 1, ενώ ανήκει στην ομάδα 2.

P(2/1) : Η πιθανότητα της λαθεμένης ταξινόμησης ενός ατόμου στην ομάδα 2, ενώ ανήκει στην ομάδα 1.

P_1 : Η εκ των προτέρων πιθανότητα ενός ατόμου να ανήκει στην ομάδα 1.

P_2 : Η εκ των προτέρων πιθανότητα ενός ατόμου να ανήκει στην ομάδα 2.

Το αναμενόμενο κόστος λαθεμένης ταξινόμησης ECM (Expected Cost of Misclassification) μπορεί να υπολογιστεί από την παρακάτω έκφραση:

$$ECM = c(1/2) \cdot P(1/2) \cdot P_2 + c(2/1) \cdot P(2/1) \cdot P_1 \qquad (3.1)$$

Ένας λογικός κανόνας ταξινόμησης είναι εκείνος που ελαχιστοποιεί το ECM. Μπορεί να αποδειχθεί[1] ότι ο κανόνας που ελαχιστοποιεί το ECM περιγράφεται ως εξής:

- Ένα άτομο με διάνυσμα διαχωριστικών χαρακτηριστικών X, ταξινομείται στην ομάδα 1 αν και μόνο αν:

$$\left[\frac{f_1(x)}{f_2(x)}\right] \geq \left[\frac{c_{(1/2)}}{c_{(2/1)}}\right]\left[\frac{P_2}{P_1}\right] \qquad (3.2)$$

διαφορετικά ταξινομείται στην ομάδα 2. Αν οι πυκνότητες πιθανότητας της μεταβλητής X ακολουθούν μια πολυκανονική κατανομή και στις δύο ομάδες, τότε η εξίσωση (3.2) απλοποιείται σημαντικά.[2]

Ένας άλλος κανόνας ταξινόμησης καταχωρεί ένα νέο άτομο στην ομάδα με την μεγαλύτερη εκ των υστέρων πιθανότητα P(1/Z) ή P(2/Z). Η εκ των υστέρων πιθανότητα με την οποία ένα άτομο με τυχαία μεταβλητή X είναι μέλος της ομάδας g, μπορεί να υπολογιστεί χρησιμοποιώντας το θεώρημα του Bayes:

[1] Βλέπε Johnson and Wichern «Applied Multivariate Statistical Analysis», Prentice-Hall International Editions, 1988, σελ. 485-490.

[2] Βλέπε Johnson and Wichern «Applied Multivariate Statistical Analysis», Prentice-Hall International Editions, 1988, σελ. 485-490.

$$P(1/X)=\frac{P_1 f_1(x)}{P_1 f_1(x)+P_2 f_2(x)}$$

$$P(2/X)=1-P(1/X) \tag{3.3}$$

Εντούτοις, ταξινομώντας ένα άτομο στην ομάδα 1 όταν *P(1/X)* > *P(2/X)* είναι ισοδύναμο με το να ελαχιστοποιήσουμε το ECM με ίσα κόστη λαθεμένης ταξινόμησης *c(1/2)* = *c(2/1)* και είναι επίσης ισοδύναμο με το να ελαχιστοποιήσουμε την συνολική πιθανότητα λαθεμένης ταξινόμησης TPM (Total Probability of Misclassification), που δίδεται από την σχέση:

$$TPM = P(1/2)\cdot P_2 + P(2/1)\cdot P_1 \tag{3.4}$$

Οι κανόνες ταξινόμησης που βασίζονται στην εξίσωση (3.2) δεν μπορούν να χρησιμοποιηθούν κατευθείαν όταν οι κατανομές πυκνότητας πιθανότητας των ομάδων $f_1(x)$ και $f_2(x)$ δεν είναι γνωστές. Επιπλέον, επειδή η μεταβλητή X μπορεί να περιέχει και ασυνεχή χαρακτηριστικά δεν μπορεί να υποτεθεί ότι οι πυκνότητες των ομάδων κατανέμονται κανονικά. Γι' αυτό τον λόγο, μια προσεγγιστική μέθοδος είναι αναγκαία για την εκτίμηση των $f_1(x)$, $f_2(x)$ και των εκ των υστέρων πιθανοτήτων P(1/X) και P(2/X). Μια απλή μέθοδος υπολογισμού αυτών των πιθανοτήτων προτείνεται από τον Tatsuoka:

$$P(g/x)=\frac{P_g\, exp\left(-x_g^2/2\right)}{P_1\, exp\left(-x_1^2/2\right)+P_2\, exp\left(-x_2^2/2\right)} \tag{3.5}$$

με x_1^2 και x_2^2 να υπολογίζονται από την σχέση:

$$X_g^2=\left(X-X_{mg}\right)^t W^{-1}\left(X-X_{mg}\right) \qquad g=1,2 \tag{3.6}$$

όπου X_{mg} είναι το διάνυσμα των μέσων τιμών των χαρακτηριστικών του δείγματος από την ομάδα g και W ο πίνακας μεταβλητότητας - συμμεταβλητότητας που υπολογίζεται σύμφωνα με την σχέση:

$$W_{ij} = \frac{1}{n-g} \sum_{g=1}^{g} \sum_{m=1}^{n_g} \left(X_{igm} - M_{ig} \right) \left(X_{jgm} - M_{jg} \right) \qquad (3.7)$$

όπου g είναι ο αριθμός των ομάδων, n_g το μέγεθος του δείγματος που προέρχεται από την ομάδα g, n είναι το μέγεθος του συνολικού δείγματος, X_{igk} είναι η τιμή του χαρακτηριστικού X_i του ατόμου k της ομάδας g, και M_{ig} είναι η μέση τιμή του χαρακτηριστικού i του δείγματος της ομάδας g.

Στην περίπτωση που οι αρχικές πιθανότητες P_1, P_2 είναι ίσες τότε η εξίσωση (3.6) μπορεί να χρησιμοποιηθεί σαν κανόνας ταξινόμησης μετρώντας ουσιαστικά τις αποστάσεις κάθε ατόμου από τις μέσες τιμές των ομάδων και ταξινομώντας το στην ομάδα με την μικρότερη απόσταση. Αυτή είναι και η πιο απλή περίπτωση ταξινόμησης γιατί δεν λαμβάνει υπόψη το κόστος και τις αρχικές πιθανότητες.

3.3 Ποσοστά και πίνακας ταξινόμησης

Ένας κοινός τρόπος για να ερευνηθούν τα αποτελέσματα της διαχωριστικής ανάλυσης, αφού έχει προσδιορισθεί η καλύτερη διαχωριστική συνάρτηση, είναι η εφαρμογή των κανόνων ταξινόμησης στα άτομα των δύο δειγμάτων και η δημιουργία του πίνακα ταξινόμησης. Ο πίνακας αυτός περιέχει τον αριθμό των σωστών και λαθεμένων ταξινομήσεων που παράγονται από τους κανόνες ταξινόμησης.

Αν ο όρος N_{ij} συμβολίζει τον αριθμό των ατόμων που ανήκουν πραγματικά στην ομάδα i και ταξινομούνται στην ομάδα j, τότε στην περίπτωση των δύο ομάδων, ο πίνακας ταξινόμησης είναι διαστάσεων 2x2. Ο πίνακας (3.1) είναι ένας πίνακας ταξινόμησης στην γενική του μορφή.

		Προβλεπόμενα		
		Ομάδα 1	*Ομάδα 2*	
Πραγματικά	*Ομάδα 1*	N_{11}	N_{12}	N_1
	Ομάδα 2	N_{21}	N_{22}	N_2
		$N_{.1}$	$N_{.2}$	N

Πίνακας 3.1

όπου τα στοιχεία της κύριας διαγωνίου συμβολίζουν τον αριθμό σωστών ταξινομήσεων και τα στοιχεία της δευτερεύουσας διαγωνίου τον αριθμό των λανθασμένων ταξινομήσεων. Επίσης, N_i, όπου i = 1,2, είναι ο πραγματικός αριθμός των ατόμων του δείγματος, ενώ $N_{.i}$ είναι ο προβλεπόμενος.

Από τον πίνακα (3.1) μπορούν να υπολογιστούν τα εξής ποσοστά σωστής ταξινόμησης:

$$\begin{aligned}(EF_1) &= (N_{11}/N_1)\cdot 100\% \\ (EF_2) &= (N_{22}/N_2)\cdot 100\% \\ (EF_t) &= \{(N_{11}+N_{22})/N\}\cdot 100\%\end{aligned} \tag{3.8}$$

Τα δύο πρώτα ποσοστά δίνουν την διαχωριστική ικανότητα που μπορεί να επιτευχθεί για κάθε ομάδα. Το τρίτο ποσοστό δείχνει την διαχωριστική ικανότητα για όλο τον πληθυσμό.

Αν ο μοναδικός τρόπος της διαχωριστικής ανάλυσης είναι να μεγιστοποιεί το συνολικό ποσοστό της σωστής ταξινόμησης, τότε Σαφώς το ποσοστό Ef_t είναι το καταλληλότερο κριτήριο. Αλλά μερικές φορές είναι πιο σημαντικό να μεγιστοποιείται είτε το EF_1 είτε το EF_2, όταν το αντίστοιχο κόστος λαθεμένης ταξινόμησης είναι πολύ υψηλό.

Κεφάλαιο

4

ΣΥΣΧΕΤΙΣΗ ΔΙΑΧΩΡΙΣΤΙΚΗΣ ΑΝΑΛΥΣΗΣ ΚΑΙ ΤΑΞΙΝΟΜΗΣΗΣ

4.1 Γενικά

Στο γενικό πλαίσιο, το βασικό πρόβλημα που πρέπει να επιλυθεί είναι η πρόβλεψη της κατάστασης ενός ατόμου σε μια δεδομένη χρονική περίοδο. Οι δύο στατιστικές τεχνικές που έχουν παρουσιαστεί μέχρι το σημείο αυτό, έχουν την δυνατότητα να επιλύσουν το πρόβλημα αυτό, χωριστά η κάθε μια.

Στα προηγούμενα κεφάλαια φάνηκε ότι η διαχωριστική ανάλυση είναι καθαρά μια τεχνική διαχωρισμού και η ταξινόμηση μια μέθοδος μέτρησης αυτού του διαχωρισμού και κατάταξης νέων ατόμων. Στην πραγματικότητα, η διαχωριστική ανάλυση προσφέρεται και για την κατάταξη των νέων ατόμων. Η κατάταξη αυτή γίνεται σύμφωνα με τον κανόνα ταξινόμησης (3.6), όπου μετρούνται μόνο οι αποστάσεις από τα κέντρα των ομάδων. Για παράδειγμα, παρατηρώντας το σχήμα (2.2), αν ένα άτομο έχει συντεταγμένες στην περιοχή αριστερά από την γραμμή C κατατάσσεται στην ομάδα 1, διαφορετικά κατατάσσεται στην ομάδα 2. Το πλεονέκτημα της κατάταξης με την διαχωριστική ανάλυση είναι ότι ελαχιστοποιείται η περιοχή επικάλυψης των δύο ομάδων, αλλά δεν λαμβάνονται υπόψη οι αρχικές πιθανότητες και τα κόστη λαθεμένης ταξινόμησης. Αυτό

αποτελεί ένα σοβαρό μειονέκτημα για την κατάταξη νέων ατόμων μέσω της διαχωριστικής ανάλυσης.

Απεναντίας, ένα είδος διαχωρισμού μπορεί να γίνει και με την τεχνική της ταξινόμησης, αν οι κανόνες της χρησιμοποιηθούν σε δείγματα των οποίων είναι γνωστή η προέλευσή τους. Σίγουρα, η ταξινόμηση είναι μια αδύνατη τεχνική στον διαχωρισμό δειγμάτων ή ομάδων, αφού αγνοείται εντελώς η βαρύτητα και η σταθερότητα του κάθε χαρακτηριστικού. Επομένως, η επιλογή του συνδυασμού των χαρακτηριστικών που επιτυγχάνει τον καλύτερο διαχωρισμό καταντά μια χρονοβόρα διαδικασία, αφού βασίζεται σε τυχαίους παράγοντες. Αναμφισβήτητα, όμως, η ταξινόμηση είναι ένα ισχυρό μεθοδολογικό εργαλείο προκειμένου να καταχωρηθούν νέα άτομα σε υπάρχουσες ομάδες, κάνοντας χρήση ρεαλιστικών κριτηρίων.

4.2 Συσχέτιση διαχωριστικής μεταβλητής και πιθανότητας ταξινόμησης

Έστω ότι ένας πληθυσμός αποτελείται από δύο ομάδες που ονομάζονται επιτυχίες και αποτυχίες και συμβολίζονται με τα γράμματα S και F αντίστοιχα. Παίρνονται δείγματα από τις δύο ομάδες και εφαρμόζονται οι στατιστικές τεχνικές που έχουν προαναφερθεί. Έστω ακόμη ότι τα άτομα που ανήκουν στην ομάδα S χαρακτηρίζονται από αρνητικές τιμές της μεταβλητής Z και τα άτομα της F από θετικές.

Έχει αποδειχθεί ότι υπό την προϋπόθεση της πολυκανονικότητας του πληθυσμού, οι προκύπτουσες τιμές της διαχωριστικής μεταβλητής Z θα ακολουθούν, θεωρητικά, την κανονική κατανομή στις δύο ομάδες. Οι κατανομές για τις δύο ομάδες είναι:

$N\,(1/2\alpha^2,\ \alpha^2)$ για τις αποτυχίες (4.1)

$N\,(-1/2\alpha^2,\ \alpha^2)$ για τις επιτυχίες. (4.2)

όπου ο πρώτος όρος της παρένθεσης παριστάνει την μέση τιμή και ο δεύτερος την μεταβλητότητα των δύο κανονικών κατανομών.

Με βάση το θεώρημα του Bayes που εκφράζεται από την εξίσωση (3.3), προκύπτει η παρακάτω σχέση:

$$P(S/Z) = \frac{P_S f(Z/S)}{P_S f(Z/S) + P_f f(Z/F)} \tag{4.3}$$

όπου f(Z/i) είναι η συνάρτηση πυκνότητας πιθανότητας της μεταβλητής Z και P_i είναι η εκ των προτέρων πιθανότητα της ομάδας i (i = S, F).

Η εξίσωση (4.3) αναπτύχθηκε από τον Ι.Λ. Μπακούρο στην διδακτορική του διατριβή, όπως φαίνεται από τις εξισώσεις (4.6) και (4.7) με τις οποίες μπορεί να εκτιμηθεί η πιθανότητα ταξινόμησης ενός ατόμου όταν δίνεται η διαχωριστική του μεταβλητή σαν γραμμική συνάρτηση των χαρακτηριστικών του. Η διαδικασία που ακολουθήθηκε προκειμένου να προκύψουν οι προαναφερθείσες εξισώσεις αναλύεται ως εξής:

➤ Η εξίσωση (4.3) μετασχηματίζεται στην μορφή:

$$P(S/Z) = \frac{1}{1 + e^{\left[ln\left(P_f/P_S\right) + ln\left(f(Z/F)/f(Z/S)\right)\right]}} \tag{4.4}$$

➤ Γνωρίζοντας ότι η συνάρτηση πυκνότητας πιθανότητας μιας κανονικής κατανομής είναι:

$$f_i(Z) = (2\pi)^{1/2} \sigma^{-1} exp\left[-\left(2\sigma^2\right)^{-1}(Z-\mu)^2\right]$$

όπου μ είναι η μέση τιμή και σ η τυπική απόκλιση, και αφού η μεταβλητή Z των δύο ομάδων ακολουθεί την κανονική κατανομή σύμφωνα με τις σχέσεις (4.1) και (4.2), ο λόγος των πυκνοτήτων f(Z/F) και f(Z/S) γίνεται:

$$\frac{f(Z/F)}{f(Z/S)} = \frac{(2\pi)^{-1/2} a^{-1} exp\left[-(2a^2)^{-1}(Z - 1/2 a^2)^2\right]}{(2\pi)^{-1/2} a^{-1} exp\left[-(2a^2)^{-1}(Z + 1/2 a^2)^2\right]}$$

$$\frac{f(Z/F)}{f(Z/S)} = exp\left[-(2a^2)^{-1}\left((Z - 1/2 a^2)^2 - (Z + 1/2 a^2)^2\right)\right]$$

$$\frac{f(Z/F)}{f(Z/S)} = exp\left[-(2a^2)^{-1} 2Z(-a^2)\right]$$

$$\frac{f(Z/F)}{f(Z/S)} = e^Z$$

$$P(S/Z) = \frac{1}{1 + e^{\left[ln\left(P_f/P_S\right) + Z\right]}} \qquad (4.5)$$

➤ Αλλά ο όρος $\ln(P_F/P_S)$ είναι σταθερός και η μεταβλητή Z προκύπτει από την γραμμική διαχωριστική συνάρτηση:

$$Z = U_0 + U_i X_i$$

επομένως η εξίσωση (4.5) μετασχηματίζεται στην παρακάτω μορφή:

$$P(S/Z) = \left[1 + exp(a + bZ)\right]^{-1} \qquad (4.6)$$

από την οποία προκύπτει η πιθανότητα ενός ατόμου να ανήκει στις επιτυχίες όταν είναι γνωστή η τιμή της διαχωριστικής του μεταβλητής Z. Επειδή ισχύει *P(F/Z) = 1 - P(S/Z)*, η παρακάτω εξίσωση δίνει την πιθανότητα του ιδίου ατόμου να ανήκει στις αποτυχίες.

$$P(S/Z) = \left[1 + \exp(-a - bZ)\right]^{-1} \qquad (4.7)$$

- Οι παράμετροι a και b είναι άγνωστοι, μπορούν όμως να εκτιμηθούν από τα ζεύγη τιμών *[Z, P(i/X)]* των δύο δειγμάτων, χρησιμοποιώντας την τεχνική της γραμμικής παλινδρόμησης. Οι τιμές των πιθανοτήτων *P(i/X)* μπορούν να υπολογιστούν από την σχέση (3.5). Αφού τα a και b είναι κοινά για τις εξισώσεις (4.6) και (4.7) μπορεί να χρησιμοποιηθεί μόνο η πρώτη για την εκτίμησή τους.

Επειδή η εκτίμησή που συνδέει την *P(S/Z)* και την *Z* είναι εκθετικής μορφής, θα μετατραπεί σε γραμμική με την εξής διαδικασία:

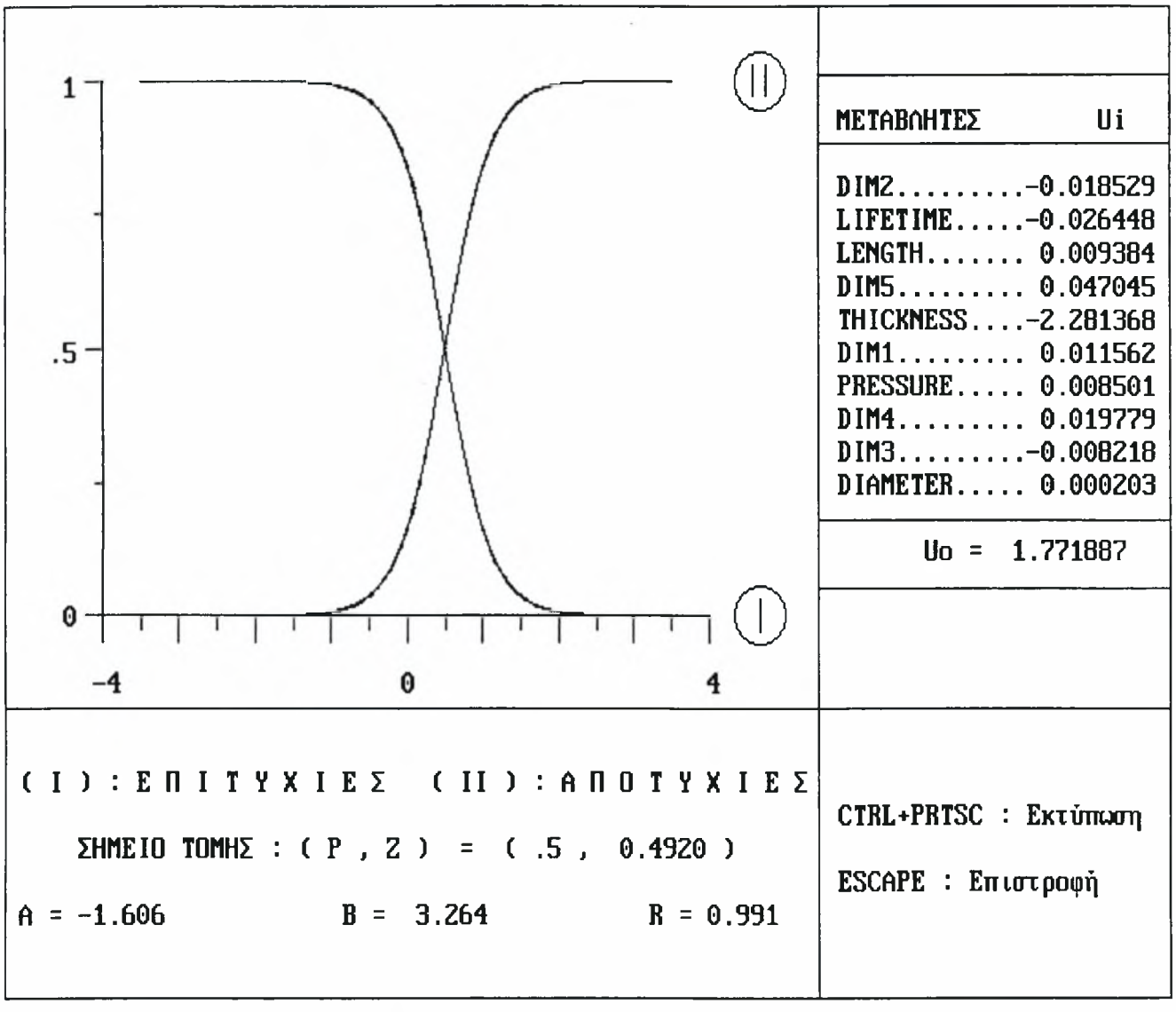
i) Αρχικά μετασχηματίζεται στην μορφή:

$$a + bZ = ln\left\{\frac{1}{P\left(S/Z\right)} - 1\right\}$$

ii) και αν τεθεί $ln\left\{\frac{1}{P\left(S/Z\right)} - 1\right\} = PR$, τότε $a + bZ = PR$.

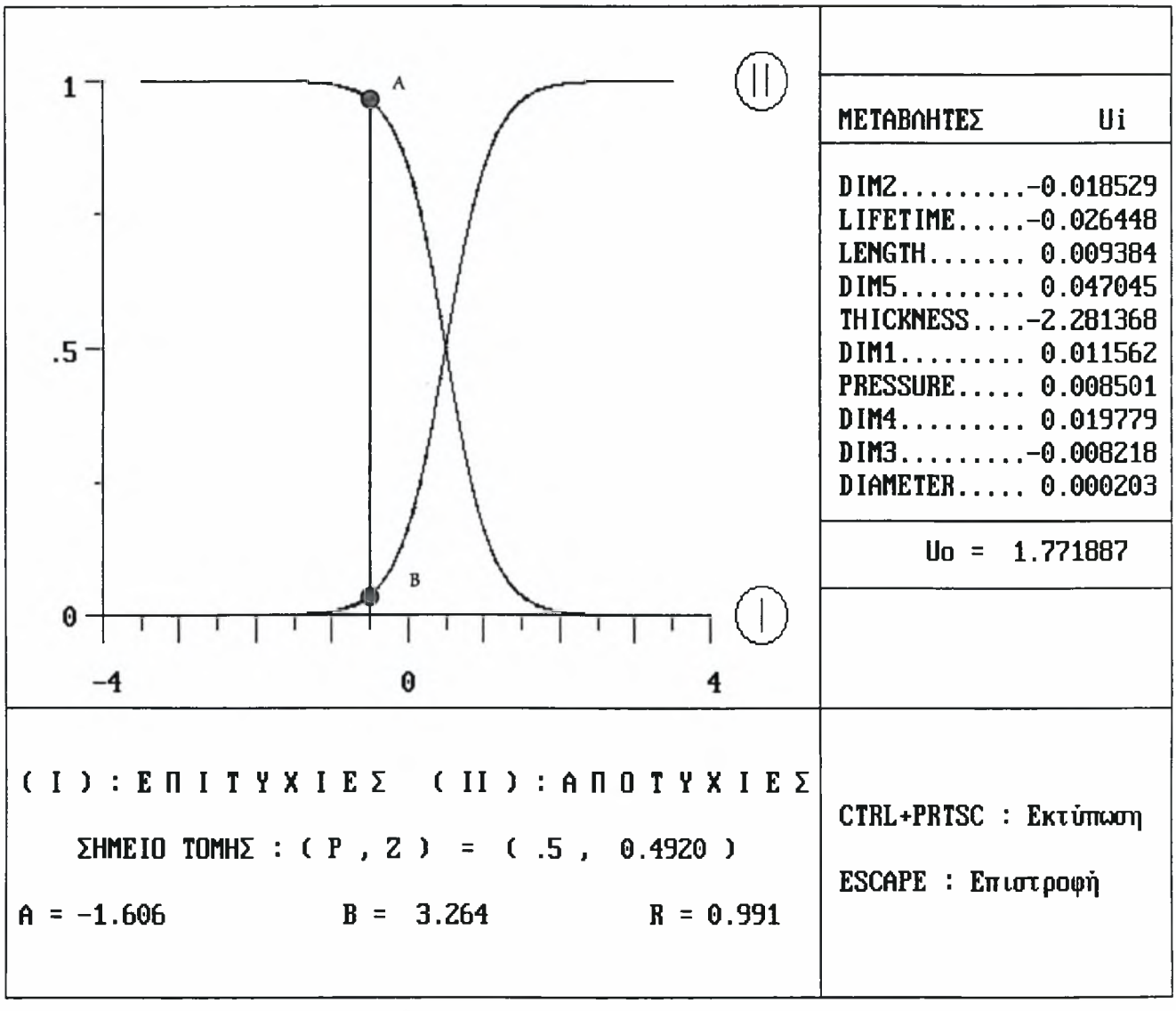
Στην συνέχεια, με την μέθοδο των ελαχίστων τετραγώνων, για τα ζεύγη τιμών *[Z, PR]*, αν στον όρο PR εισαχθούν οι *P(S/X)* αντί των *P(S/Z)*, υπολογίζονται οι αμερόληπτες εκτιμήτριες a_0 και b_0.

Ενδιαφέρον παρουσιάζει η γραφική απεικόνιση των εξισώσεων (4.6) και (4.7), που παριστάνεται στο σχήμα (4.1). Εκεί, φαίνεται η μεταβολή που υφίσταται η πιθανότητα ενός ατόμου να ανήκει σε μια ομάδα καθώς αυξάνει η τιμή της διαχωριστικής μεταβλητής. Αξιοσημείωτο είναι το γεγονός ότι το σχήμα (4.1) ισχύει για όλο τον πληθυσμό επειδή οι παράμετροι a_0 και b_0 υπερβαίνουν τα στενά όρια του συγκεκριμένου δείγματος από το οποίο εκτιμήθηκαν. Έτσι, βάσει του σχήματος (4.1) μπορεί να προβλεφθεί η τελική κατάσταση ενός νέου ατόμου, χωρίς να είναι γνωστή η θεωρία της διαχωριστικής ανάλυσης και της ταξινόμησης. Επιπλέον, η πρόβλεψη αυτή όχι μόνο δεν είναι εμπειρική αλλά στηρίζεται σε θεμελιωμένες στατιστικές τεχνικές.



Σχήμα 4.1

Διάγραμμα πιθανότητας ταξινόμησης - διαχωριστικής μεταβλητής



Σχήμα 4.2

Η διαδικασία με την οποία προβλέπεται η τελική κατάσταση ενός ατόμου, είναι η εξής:

- Μετρούνται τα χαρακτηριστικά του νέου ατόμου.
- Εισάγονται στην διαχωριστική συνάρτηση , μαζί με τους συντελεστές που υπάρχουν στο διάγραμμα και υπολογίζεται η διαχωριστική μεταβλητή.
- Υψώνεται ευθεία κάθετη στον άξονα Z από την τιμή που υπολογίστηκε, η οποία τέμνει τις καμπύλες P(S/Z) και P(F/Z) στα σημεία A και B.
- Αν το σημείο A είναι πάνω από το σημείο B, τότε η πιθανότητα του ατόμου να ανήκει στις επιτυχίες είναι μεγαλύτερη από την πιθανότητα να ανήκει στις αποτυχίες, και επομένως κατατάσσεται στην ομάδα των επιτυχιών. Στην αντίθετη περίπτωση κατατάσσεται στην ομάδα των αποτυχιών.

Σε τελική ανάλυση, κάνοντας χρήση του διαγράμματος ταξινόμησης μπορούμε να καταλήξουμε σε συμπέρασμα για την κατάταξη ενός νέου ατόμου, αποφεύγοντας την χρήση περίπλοκων και χρονοβόρων στατιστικών εννοιών.

4.3 Έλεγχος εφαρμογής του διαγράμματος ταξινόμησης

Όταν μέσα στην διαχωριστική μεταβλητή υπάρχουν χαρακτηριστικά που δεν κατανέμονται κανονικά ή τα οποία είναι ασυνεχή, η υπόθεση της πολυκανονικότητας είναι αμφισβητήσιμη και η παραπάνω θεωρία δεν μπορεί να εφαρμοστεί άμεσα. Στην περίπτωση αυτή πρέπει να γίνει ένας έλεγχος προσαρμογής κανονικής κατανομής της μεταβλητής Z και των παραμέτρων της.

Αν ο έλεγχος δείξει ότι η κατανομή της μεταβλητής Z αποκλίνει από την ιδεατή, για κάποια από τις δύο ομάδες, τότε για την ικανοποιητική εφαρμογή του διαγράμματος σε όλο τον πληθυσμό, πρέπει ο βαθμός προσαρμογής της γραμμής παλινδρόμησης $a_0 + b_0 = PR$ στα ζεύγη τιμών [Z, PR] του δείγματος να είναι υψηλός και στην συνέχεια να ελεγχθεί αν η σχέση των Z και PR είναι προσεγγιστικά γραμμική σε όλο τον πληθυσμό.

4.4 Έλεγχος κανονικότητας της διαχωριστικής μεταβλητής

Ο έλεγχος της κανονικότητας της κατανομής γίνεται χρήσει της μεθόδου Kolmogorov - Smirnov[1]. Ο έλεγχος αυτός μπορεί να εφαρμοσθεί με οποιοδήποτε μέγεθος δείγματος και ενδείκνυται στις περιπτώσεις που το μέγεθος του δείγματος είναι μικρό.

Αρχικά, με την μέθοδο της μέγιστης πιθανότητας εκτιμούνται οι παράμετροι της κανονικής κατανομής. Κατόπιν, με τις εκτιμήτριες που έχουν υπολογιστεί προσδιορίζονται οι αθροιστικές πιθανότητες Φ(Z) από τους πίνακες της κανονικής κατανομής.

Στην συνέχεια ταξινομούνται οι τιμές της μεταβλητής Z του δείγματος κατά αύξουσα τάξη. Σχηματίζεται η αθροιστική συνάρτηση συχνότητας Φ*(Z) που αρχίζει από τον μηδέν και αυξάνεται κλιμακωτά, με βήμα $1/n_g$, μέχρι την μονάδα. Όπου n_g είναι το μέγεθος του δείγματος που αντιπροσωπεύει την ομάδα g.

Τέλος, υπολογίζεται η μέγιστη διαφορά:

$$D_n(Z) = max\left[\Phi^*(Z) - \Phi(Z)\right]$$

Αν η τιμή της $D_n(Z)$ είναι μεγαλύτερη ή ίση από το D_Z, η υπόθεση της κανονικότητας της Z για την συγκεκριμένη ομάδα απορρίπτεται. Η τιμή του D_Z διαβάζεται από τους πίνακες Kolmogorov - Smirnov για ορισμένο επίπεδο σημαντικότητας και μέγεθος δείγματος.

4.5 Συντελεστής συσχέτισης δείγματος και πληθυσμού

Για να μετρηθεί ο βαθμός προσαρμογής της γραμμής παλινδρόμησης στα ζεύγη τιμών [Z, PR] του δείγματος, είναι αναγκαίο να υπολογιστεί ο συντελεστής συσχέτισης R για το δείγμα, σύμφωνα με την ακόλουθη σχέση:

[1] Βλέπε Δ.Π. Ψωϊνός, «Εφαρμοσμένη Στατιστική», σελ. 237, εκδόσεις Ζήτη, 1992.

$$R = \frac{n\sum_{i=1}^{n} Z_i PR_i - \left(\sum_{i=1}^{n} Z_i\right)\left(\sum_{i=1}^{n} PR_i\right)}{\sqrt{\left[n\sum_{i=1}^{n} Z_i^2 - \left(\sum_{i=1}^{n} Z_i\right)^2\right]\left[n\sum_{i=1}^{n} PR_i^2 - \left(\sum_{i=1}^{n} PR_i\right)^2\right]}}$$

Όσο πλησιέστερα προς την μονάδα βρίσκεται η τιμή του συντελεστή συσχέτισης, τόσο καλύτερη είναι η προσαρμογή της γραμμής παλινδρόμησης στα ζεύγη [Z, PR] του δείγματος.

Ακόμη και αν η τιμή του συντελεστή συσχέτισης για το δείγμα προσεγγίζει την μονάδα, τίθεται το ερώτημα αν το διάγραμμα μπορεί να εφαρμοσθεί για όλο τον πληθυσμό. Στην περίπτωση αυτή πρέπει να γίνει ένας έλεγχος υποθέσεων για την τιμή του συντελεστή συσχέτισης του πληθυσμού. Οι υποθέσεις που πρέπει να διατυπωθούν είναι οι εξής:

- Μηδενική υπόθεση (H_0) : Η τιμή του συντελεστή συσχέτισης για τον πληθυσμό είναι $\rho = 0$.
- Εναλλακτική υπόθεση (H_1) : Η τιμή του συντελεστή είναι $\rho > 0$.

Έχει αποδειχθεί ότι όταν $\rho = 0$, τότε η μεταβλητή t ακολουθεί την κατανομή Student με (n - 2) βαθμούς ελευθερίας, όπου t:

$$t = \frac{R}{\sqrt{\frac{1-R^2}{n-2}}}$$

και n και R το μέγεθος και ο συντελεστής συσχέτισης του δείγματος αντίστοιχα.

Όταν η τιμή της t που υπολογίζεται είναι μεγαλύτερη από εκείνη που προκύπτει από τους πίνακες της κατανομής Student, τότε η μηδενική υπόθεση H_0 απορρίπτεται, για το συγκεκριμένο μέγεθος δείγματος και επίπεδο σημαντικότητας που έχει επιλεχθεί.

<u>Παρατήρηση:</u> Οι πίνακες της κατανομής Student παρέχουν τιμές της μεταβλητής t αμφίπλευρα για α/2, με επίπεδο σημαντικότητας α. Δεδομένου ότι η εναλλακτική υπόθεση H_1 είναι μονόπλευρη, πρέπει να υπολογίζουμε την τιμή του t από τους πίνακες για επίπεδο σημαντικότητας 2α.

Συμπερασματικά, στην περίπτωση απόρριψης της μηδενικής υπόθεσης H_0, η σχέση ανάμεσα στις μεταβλητές Z και PR μπορεί να θεωρηθεί κατά προσέγγιση γραμμική. Αυτό σημαίνει ότι η γραμμή παλινδρόμησης $a_0 + b_0Z = PR$, μπορεί να χρησιμοποιηθεί με ικανοποιητική ακρίβεια για όλο τον πληθυσμό. Άρα, η χρήση του διαγράμματος ενδείκνυται για την πρόβλεψη ενός οποιουδήποτε νέου ατόμου του αν θα ανήκει στις επιτυχίες ή τις αποτυχίες.

Κεφάλαιο

5

ΧΡΗΣΗ ΟΜΑΔΟΠΟΙΗΜΕΝΩΝ ΜΕΤΑΒΛΗΤΩΝ ΓΙΑ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗ ΠΡΟΒΛΕΨΗΣ ΑΞΙΟΠΙΣΤΙΑΣ

5.1 Γενικά

Σκοπός των κεφαλαίων που ακολουθούν είναι η μελέτη της αξιοπιστίας των υποθαλάσσιων αγωγών πετρελαίου που είναι εγκατεστημένοι στην Βόρεια Θάλασσα. Για τον σκοπό αυτό, γίνεται χρήση των πολυμεταβλητών στατιστικών τεχνικών της διαχωριστικής ανάλυσης και της ταξινόμησης. Οι βασικές αρχές των δύο αυτών τεχνικών αναπτύχθηκαν στα προηγούμενα κεφάλαια.

Η ανάλυση ξεκινά με την μελέτη των δεδομένων των αγωγών και τον διαχωρισμό τους σε δύο ομάδες: των επιτυχιών και των αποτυχιών. Στην ομάδα των επιτυχιών περιλαμβάνονται εκείνοι οι αγωγοί οι οποίοι καθ' όλη την διάρκεια της συλλογής των δεδομένων δεν παρουσίασαν καμία βλάβη. Αντίστοιχα στην ομάδα των αποτυχιών περιλαμβάνονται όλοι εκείνοι οι αγωγοί οι οποίοι παρουσίασαν τουλάχιστον μία φορά βλάβη, είτε πριν είτε μετά την τοποθέτηση και λειτουργία τους.

Εν συνεχεία, γίνεται η επιλογή εκείνων των χαρακτηριστικών των αγωγών (ή ενδεχομένως και άλλων παραμέτρων που μπορεί να απαιτούνται), που θα χρησιμοποιηθούν στην ανάλυση. Αντικειμενικός σκοπός η επιλογή πλέον εκεί-

νων των χαρακτηριστικών τα οποία θα παρουσιάζουν την μεγαλύτερη διαχωριστική ισχύ και θα παρέχουν εκείνη την διαχωριστική συνάρτηση που θα περιγράφει με τον βέλτιστο μαθηματικό τρόπο την συμπεριφορά των αγωγών.

Η οργάνωση για την επιλογή των χαρακτηριστικών αυτών γίνεται με την δημιουργία διαφόρων σεναρίων, τα οποία περιλαμβάνουν διαφορετικούς συνδυασμούς χαρακτηριστικών. Κάθε ένα από τα σενάρια αυτά αναλύεται χωριστά έτσι ώστε τελικά να παρουσιάζεται μια γενική συγκριτική εικόνα αυτών.

Τέλος, με τον συνδυασμό των δύο τεχνικών της διαχωριστικής ανάλυσης και της ταξινόμησης, γίνεται η εκτίμηση της πιθανότητας ενός αγωγού, εκ των προτέρων, να ανήκει στην ομάδα των επιτυχιών ή των αποτυχιών, απλώς έχοντας μετρήσει τα χαρακτηριστικά του αγωγού τα οποία έχουν προεπιλεγεί.

Βασική ιδιαιτερότητα στην ανάλυση που ακολουθεί αποτελεί η εισαγωγή ομαδοποιημένων χαρακτηριστικών, πέρα των απλών μεταβλητών. Οι τιμές των χαρακτηριστικών αυτών δεν παρέχονται αυτούσιες από τα δεδομένα, αλλά αποτελούν συνδυασμούς άλλων χαρακτηριστικών.

Σημειώνεται ότι τα στοιχεία των αγωγών πετρελαίου για την περιοχή της Βορείου Θάλασσας, παραχωρήθηκαν από την εταιρία Battelle της Ελβετίας για εκπαιδευτικούς σκοπούς και για τον λόγο αυτό δεν δημοσιεύονται.

5.2 Γενική παρουσίαση σεναρίων

Στο σημείο αυτό, γίνεται μια σύντομη περιγραφή της δομής των διαφόρων σεναρίων, όπως αυτά παρουσιάζονται στις επόμενες ενότητες μαζί με τους πίνακες και τα διαγράμματα.

Αρχικά, αφού επιλεγούν τα χαρακτηριστικά - μεταβλητές τα οποία θα επεξεργαστούν σε κάθε περίπτωση, δημιουργούνται οι βάσεις δεδομένων των επιτυχιών και των αποτυχιών.

Στον πρώτο πίνακα παρουσιάζονται τα συνολικά στατιστικά στοιχεία που προέρχονται από την ανάλυση των δεδομένων των δύο βάσεων. Έτσι, υπολογίζονται η συνολική μέση τιμή και τυπική απόκλιση καθώς και η μέγιστη και ελάχιστη τιμή του κάθε χαρακτηριστικού.

Ο δεύτερος πίνακας παρουσιάζει τα στοιχεία της διαχωριστικής ανάλυσης. Για κάθε χαρακτηριστικό εμφανίζονται οι συντελεστές της διαχωριστικής συνάρτησης U_i, οι κανονικοποιημένοι συντελεστές S_i και επίσης προσδιορίζεται η βαρύτητα και η πιθανή ευστάθεια ή αστάθεια του κάθε χαρακτηριστικού. Ο σταθερός όρος της διαχωριστικής συνάρτησης U_0 εμφανίζεται στην τελευταία στήλη. Με βάση τα στοιχεία αυτά, υπολογίζεται η τιμή της διαχωριστικής μεταβλητής Z_m, η οποία προκύπτει από την ακόλουθη γραμμική διαχωριστική συνάρτηση:

$$Z_m = U_0 + U_1 X_{1m} + U_2 X_{2m} + ... + U_i X_{im}$$

Στην συνέχεια εφαρμόζεται η θεωρία της ταξινόμησης στα άτομα που αποτελούν τις δύο βάσεις δεδομένων. Αφού ορισθεί η εκ των προτέρων πιθανότητα επιτυχίας, εμφανίζεται ο πίνακας ταξινόμησης. Επίσης, παρουσιάζονται τα ποσοστά σωστής ταξινόμησης για τις επιτυχίες EF_1, για τις αποτυχίες EF_2 και για το σύνολο EF_t.

Ακολουθεί το διάγραμμα ταξινόμησης, το οποίο συσχετίζει την διαχωριστική μεταβλητή Z με την εκ των υστέρων πιθανότητα ταξινόμησης. Στον οριζόντιο άξονα βρίσκεται η μεταβλητή Z, ενώ στον κάθετο άξονα οι υπό συνθήκη πιθανότητες. Εμφανίζονται ακόμη οι συντελεστές των χαρακτηριστικών της διαχωριστικής συνάρτησης U_i, καθώς και ο σταθερός όρος U_0. Δίδονται επίσης οι συντελεστές a και b της εξίσωσης παλινδρόμησης για τα ζεύγη τιμών [Z, PR] και ο συντελεστής R, ο οποίος προκύπτει από τα ζεύγη αυτά. Ο συντελεστής R αποτελεί, πρακτικά, ένα μέτρο της προσέγγισης που επιτυγχάνεται από την εξίσωση παλινδρόμησης.

Ακολουθούν κάποιοι στατιστικοί έλεγχοι, βάσει των οποίων ελέγχεται η ποιότητα των μετρήσεων που διαθέτουμε και αν τα συμπεράσματα που προκύπτουν από τα συγκεκριμένα δείγματα μπορούν να ισχύσουν για ολόκληρο τον πληθυσμό.

- **Έλεγχος μέσων τιμών.**

Όταν η τιμή του Wilk's Λ προσεγγίζει το μηδέν, ο διαχωρισμός ανάμεσα στις δύο ομάδες είναι αρκετά καλός. Επίσης, δεδομένης της τιμής του F, και με

βάση την προσέγγιση κατά Rao, από τους πίνακες της κατανομής F (για v_1, v_2 βαθμούς ελευθερίας και επίπεδο σημαντικότητας α%), υπολογίζουμε την τιμή F_{v_1,v_2}. Αν ισχύει $F_{v_1,v_2} < F$, τότε η τιμή του Wilk's Λ ισχύει για ολόκληρο τον πληθυσμό.

- **Έλεγχος ισότητας διασπορών.**

Ο έλεγχος γίνεται υπολογίζοντας την παράμετρο Box's B και την τιμή της μεταβλητής F για v_1 και v_2 βαθμούς ελευθερίας.

- **Έλεγχος συσχέτισης.**

Εξετάζεται κατά πόσο συσχετίζεται η υπολογισθείσα διαχωριστική συνάρτηση με τις μετρήσεις των χαρακτηριστικών. Ελέγχοντας τις ιδιοτιμές του πίνακα $W^{-1}B$ και τις τιμές του Κανονικού Συντελεστή Συσχέτισης (C.C.C.), διακρίνεται η ιδιοτιμή από την οποία προέκυψε η διαχωριστική συνάρτηση, ενώ ο συντελεστής συσχέτισης μας παρέχει πληροφορίες όσον αφορά το κατά πόσο η διαχωριστική συνάρτηση περιγράφει επαρκώς τις διαφορές ανάμεσα στις δύο ομάδες. Όπως έχει ήδη αναφερθεί, τιμές του C.C.C. που προσεγγίζουν την μονάδα υποδεικνύουν υψηλή συσχέτιση ανάμεσα στην διαχωριστική συνάρτηση και τις ομάδες.

- **Έλεγχος κανονικότητας της μεταβλητής Z.**

Εφαρμόζεται ο έλεγχος προσαρμογής της κανονικής κατανομής Kolmogorov - Smirnov, για επίπεδα σημαντικότητας 1%, 5%, 10% και υπολογίζονται οι εκτιμήτριες της μέσης τιμής και της μεταβλητότητας των δύο ομάδων.

- **Έλεγχος παλινδρόμησης.**

Ελέγχεται η ύπαρξη συσχέτισης μεταξύ των τιμών της μεταβλητής Z και των εκ των υστέρων πιθανοτήτων, για ολόκληρο τον πληθυσμό. Γίνεται ο σχετικός έλεγχος t-Student για επίπεδο σημαντικότητας 5% και παρατηρείται κατά πόσο το διάγραμμα ταξινόμησης ενδείκνυται για την πρόβλεψη οποιουδήποτε νέου ατόμου να ανήκει στην ομάδα των επιτυχιών ή των αποτυχιών.

5.3 Η σημασία της χρήσης ομαδοποιημένων μεταβλητών στην αξιοπιστία

Οι πολυμεταβλητές στατιστικές τεχνικές της διαχωριστικής ανάλυσης και της ταξινόμησης επεξεργάζονται τις τιμές διαφόρων χαρακτηριστικών με σκοπό την αλγεβρική διατύπωση και περιγραφή των διαφορών των ατόμων των εξεταζομένων ομάδων. Η συμπεριφορά των ατόμων έχει άμεση σχέση σε κάθε περίπτωση με τις τιμές των χαρακτηριστικών που επιλέγονται και τα οποία φέρουν τον «χαρακτήρα» και τις ιδιότητες του κάθε ατόμου. Έτσι, γίνεται σαφής η σπουδαιότητα της επιλογής των χαρακτηριστικών στα στατιστικά αποτελέσματα και την εγκυρότητα αυτών.

Στην παρούσα μελέτη, εκτός από την χρησιμοποίηση των απλών χαρακτηριστικών, επιχειρείται η εισαγωγή και ομαδοποιημένων μεταβλητών. Η τιμές αυτών προκύπτουν από συνδυασμούς άλλων απλών χαρακτηριστικών.

Η σημασία της χρήσης τέτοιου είδους μεταβλητών αποδεικνύεται ιδιαίτερα μεγάλη και ταυτόχρονα πολύπλευρη. Βασικό χαρακτηριστικό των ομαδοποιημένων μεταβλητών είναι το ότι «φέρουν» μεγαλύτερο πλήθος πληροφοριών για τα άτομα των ομάδων, καθώς αποτελούν συνδυασμούς άλλων μεταβλητών.

Παράλληλα όμως, μελετάται και η αλληλεπίδραση ανάμεσα σε χαρακτηριστικά, τα οποία ως μεμονωμένα, ενδεχομένως, δεν αποτελούν ικανό διαχωριστικό μέσο. Έτσι, για παράδειγμα, η πρακτική αποδεικνύει ότι ένας αγωγός μεγάλου μήκους με μικρό χρόνο ζωής συμπεριφέρεται διαφορετικά απ' ότι έναν αγωγό μικρού μήκους και μεγάλης διάρκειας ζωής. Συνδυάζοντας, λοιπόν, τα χαρακτηριστικά «Μήκος» και «Χρόνο ζωής» ενός αγωγού, δημιουργείται ένα νέο χαρακτηριστικό που περιγράφει την ιδιότητα αυτή.

Συν τοις άλλοις, η ταυτόχρονη παρουσία ενός χαρακτηριστικού μεμονωμένα καθώς και μέσα σε ομαδοποιημένες μεταβλητές, ενδέχεται να αναβαθμίσει ή να υποβαθμίσει την διαχωριστική ισχύ της συνάρτησης, ανάλογα με τον αν «τονίζεται» η ισχύς του ή αν οι πληροφορίες που «φέρει» καθίστανται πλεονάζουσες, αντίστοιχα.

Μια άλλη σημαντική ιδιαιτερότητα των ομαδοποιημένων χαρακτηριστικών, αποτελεί το γεγονός ότι μπορούν να αποτελούν, εκτός από απλούς συνδυα-

σμούς, και παραστάσεις οι οποίες εκφράζουν την φυσική σημασία ενός φαινομένου π.χ. ή μιας ιδιότητας. Η ικανότητα αυτή των ομαδοποιημένων μεταβλητών είναι ιδιαίτερα σημαντική δεδομένου ότι πλέον λαμβάνουμε υπόψη και την βαρύτητα φυσικών μεγεθών και παραγόντων που μπορούν να επηρεάζουν σε μεγάλο βαθμό, ενδεχομένως, την συμπεριφορά και την αξιοπιστία ενός αγωγού.

Η παραπάνω διαπίστωση, λαμβάνει μεγαλύτερη σημασία στην περίπτωση της χρησιμοποίησης **αδιάστατων** ομαδοποιημένων χαρακτηριστικών. Η χρήση αδιάστατων μεγεθών πλεονεκτεί κυρίως στο ότι επιτυγχάνεται η απαλοιφή των μονάδων του κάθε χαρακτηριστικού, ανεξαρτητοποιώντας ουσιαστικά την ανάλυση από τις διαστατικές διαφορές μεταξύ των μεγεθών. Επιπρόσθετα, οι αδιάστατες μεταβλητές έχουν άμεση σχέση με την ερμηνεία διαφόρων φαινομένων που μπορεί να λαμβάνουν χώρα στο πεδίο της μελέτης. Έτσι, π.χ. η επιλογή ως χαρακτηριστικού της αδιάστατης ομαδοποιημένης μεταβλητής Re (αριθμός Reynolds), συμπεριλαμβάνει στην ανάλυση την επίδραση των ροϊκών φαινομένων που λαμβάνουν χώρα στο εσωτερικό ενός αγωγού.

Συμπερασματικά, με την εισαγωγή ομαδοποιημένων μεταβλητών στο σύστημα επιτυγχάνεται η αποτελεσματικότερη και ακριβέστερη πρόβλεψη αξιοπιστίας, συνδυάζοντας παράλληλα την ανάλυση αυτή με άλλες παραμέτρους (Μηχανική των Ρευστών, Μετάδοση Θερμότητας, Δυναμική Κυμάτων κ.ά.) που ενδεχομένως έχουν άμεση επίδραση στην συμπεριφορά του συστήματος.

5.4 Σενάρια μελέτης αξιοπιστίας των αγωγών της βορείου θάλασσας

Όλες οι πληροφορίες που διατίθονται για την περαιτέρω μελέτη, δομούνται σε δύο κύριες βάσεις δεδομένων, τα inventory files (αρχεία απογραφής) και τα incident files (αρχεία γεγονότων). Τα δεδομένα χρονολογούνται από τις αρχές του έτους 1967 μέχρι περίπου τα τέλη του 1986. Στην πρώτη κατηγορία συμπεριλαμβάνονται τα ακόλουθα στοιχεία για όλους τους αγωγούς:

1. Ένας κωδικός αριθμός μητρώου για κάθε αγωγό,
2. Την χώρα στην οποία καταλήγει ο αγωγός,
3. Η περιοχή από την οποία ξεκινά ο αγωγός,
4. Η διαδρομή που ακολουθεί ο αγωγός,
5. Η ιδιοκτήτρια εταιρία του αγωγού,
6. Το προϊόν που μεταφέρει ο αγωγός (πετρέλαιο, φυσικό αέριο κλπ.),
7. Η ημερομηνία εγκατάστασης του αγωγού,
8. Το μήκος του αγωγού (χιλιόμετρα),
9. Η διάμετρος του αγωγού (inches),
10. Το πάχος του τοιχώματος του αγωγού (inches),
11. Το υλικό κατασκευής του αγωγού,
12. Το χρησιμοποιούμενο αντιδιαβρωτικό και αντιοξειδωτικό υλικό,
13. Την χρησιμοποιούμενη καθοδική προστασία,
14. Το είδος επικάλυψης του αγωγού,
15. Το πάχος της επικάλυψης του αγωγού,
16. Η πίεση λειτουργίας του αγωγού (bar),
17. Πόσες φορές έχει αποτύχει ο αγωγός,
18. Οι ημερομηνίες τις οποίες παρουσιάστηκαν οι αποτυχίες.

Στην δεύτερη κατηγορία, περιλαμβάνονται οι ακόλουθες πληροφορίες για τους αγωγούς που αστόχησαν:

1. Ο κωδικός αριθμός του κατάλογου συμβάντων,
2. Ο κωδικός αριθμός του αγωγού, του μητρώου απογραφών,
3. Η διαδρομή του αγωγού,
4. Η ημερομηνία εγκατάστασης του αγωγού,
5. Η ημερομηνία όπου έλαβε χώρα η αποτυχία,
6. Η κατάσταση λειτουργίας όταν συνέβη η αποτυχία (hydrotest, εγκατάσταση, λειτουργία κλπ.),
7. Η τοποθεσία όπου συνέβη η αστοχία,
8. Το βάθος του σημείου όπου παρουσιάστηκε η βλάβη (μέτρα),
9. Το γεγονός της αποτυχίας του αγωγού,

10. Το αίτιο της αποτυχίας του αγωγού,
11. Η μέθοδος επισκευής του αγωγού.

Για την συγκεκριμένη ανάλυση, λήφθηκαν υπόψη τα ακόλουθα χαρακτηριστικά:

Χαρακτηριστικό	Μονάδες
Μήκος αγωγού	(Km)
Διάμετρος του αγωγού	(inches)
Πάχος του τοιχώματος	(inches)
Πίεση λειτουργίας	(bar)
Υλικό κατασκευής-όριο διαρροής	(N/mm^2)
Χρόνος Ζωής	(months)

Επίσης, δημιουργήθηκαν οι ακόλουθες ομαδοποιημένες μεταβλητές, από τις οποίες οι τέσσερις (4) είναι αδιάστατες:

Μεταβλητή	Τύπος	Μονάδες
DIM1	$(Grade)/0{,}1\cdot(Pressure)$	[-]
DIM2	$(Length)/2{,}54\cdot(Diameter)\times 100$	[-]
DIM3	$(Thickness)/(Diameter)\times 1000$	[-]
DIM4	$\frac{\sigma_L\cdot\sigma_E\cdot 0{,}1^2}{(Grade)^2}\times 100$	[-]
DIM5	$\frac{2{,}54\cdot 10^{-2}\pi}{10^3}(Diameter)(Length)(Life-time)$	[$m^2\cdot$months]

- Η μεταβλητή DIM1 αποτελεί ουσιαστικά έναν συντελεστή ασφαλείας λειτουργίας του αγωγού, όσον αφορά την *πίεση λειτουργίας* του σε σχέση με την *τάση διαρροής* του υλικού.
- Η μεταβλητή DIM2 εκφράζει τον αδιάστατο λόγο του *μήκους* του αγωγού ως προς την *διάμετρό* του. Η μεταβλητή αυτή παρουσιάζει ιδιαίτερο ενδιαφέρον, δεδομένου ότι συνδυάζει δύο χαρακτηριστικά τα οποία συνδυαστικά επηρεάζουν σαφώς την αξιοπιστία του αγωγού. Έτσι, π.χ. ένας αγωγός μεγάλου μήκους θα πρέπει να έχει μεγαλύτερη διάμετρο από έναν αγωγό μικρότερου μήκους, βάσει της απλής πρακτικής. Λόγω των μεγάλων μεγεθών που προκύπτουν, οι τελικές τιμές έχουν διαιρεθεί με το 1000.
- Η μεταβλητή DIM3 εισάγει τον αδιάστατο λόγο του *πάχους* των τοιχωμάτων του αγωγού και της *διαμέτρου* του. Λόγω των μικρών μεγεθών που προκύπτουν, οι τελικές τιμές έχουν πολλαπλασιασθεί με το 1000.
- Η μεταβλητή DIM4 αποτελεί ουσιαστικά έναν συνδυασμό τριών παραμέτρων:
  - της τάσης διαρροής του υλικού του αγωγού (Grade),
  - της εφελκυστικής τάσης κατά μήκος του σωλήνα $\sigma_L = \frac{P \cdot D}{4T}$,
  - και της εφαπτομενικής τάσης $\sigma_E = \frac{P \cdot D}{2T}$.

  Λόγω των μικρών μεγεθών που προκύπτουν, οι τελικές τιμές έχουν πολλαπλασιασθεί με το 100.
- Η μεταβλητή DIM5 εκφράζει τον λόγο της επιφάνειας του αγωγού επί τον χρόνο ζωής του.

Σε μια πρώτη ανάλυση που προηγήθηκε, η οποία συμπεριλάμβανε όλα τα απλά χαρακτηριστικά που αναφέρθηκαν, καθώς επίσης και την μεταβλητή product (προϊόν), οι μεταβλητές Grade και Product παρουσίασαν την μικρότερη διαχωριστική ισχύ, και για τον λόγο αυτό δεν συμπεριλήφθηκαν στην περαιτέρω μελέτη, πλην του χαρακτηριστικού Grade το οποίο εμφανίζεται στην ομαδοποιημένη μεταβλητή DIM5. Αυτό έγινε λόγω του περιορισμού του προγράμματος του Η/Υ στον αριθμό των χρησιμοποιούμενων χαρακτηριστικών σε δέκα (10).

Σημειώνεται ότι υπήρξαν περιπτώσεις αγωγών οι οποίοι αστόχησαν περισσότερες από μία φορές. Κάθε επιπρόσθετη αποτυχία προσμετρήθηκε στην ομάδα των αποτυχιών κανονικά. Επίσης, αγωγοί οι οποίοι αστόχησαν πριν ή κατά την διάρκεια της εγκατάστασης, θεωρούνται ως αποτυχόντες και συγκαταλέγονται στην ομάδα των αποτυχιών, με μηδενικό όμως χρόνο ζωής.

Τέλος, αναφέρεται ότι το αρχείο απογραφής (inventory file) περιλαμβάνει συνολικά 138 αγωγούς από τους οποίους οι 29 αστόχησαν, σημειώνοντας 59 αποτυχίες, και οι υπόλοιποι 109 δεν παρουσίασαν καμία βλάβη.

ΕΝΟΤΗΤΑ Β

Πρακτικη Εφαρμογη Τεχνικων Διαχωριστικης Αναλυσης και Ταξινομησης

Κεφάλαιο

6

ΜΕΛΕΤΗ ΑΞΙΟΠΙΣΤΙΑΣ ΑΓΩΓΩΝ ΠΕΤΡΕΛΑΙΟΥ ΤΗΣ ΒΟΡΕΙΟΥ ΘΑΛΑΣΣΑΣ

6.1 Γενικά

Στις ακόλουθες παραγράφους, παρουσιάζεται λεπτομερειακά η ανάλυση της αξιοπιστίας των αγωγών πετρελαίου στην Βόρεια Θάλασσα, βάσει των τεχνικών της διαχωριστικής ανάλυσης και της ταξινόμησης. Η πρακτική μελέτη της αξιοπιστίας, περιλαμβάνει την παρουσίαση και ανάλυση 22 σεναρίων τα οποία αποτελούνται από συνδυασμούς 10 διαφορετικών μεταβλητών, ομαδοποιημένων και μη.

Τα βήματα που ακολουθούνται σε κάθε περίπτωση είναι αυτά που ακριβώς περιγράφηκαν στην παράγραφο §5.2. Σε κάθε σημείο αποτυπώνονται και σχολιάζονται τα, κατά βήμα, αποτελέσματα που εξάγει το πρόγραμμα του Η/Υ.

Απώτερος σκοπός της διερεύνησης αυτής, η αναζήτηση εκείνου του σεναρίου που στο σύνολό του εμφανίζει τις βέλτιστες στατιστικές παραμέτρους που μελετούμε και την μεγαλύτερη ικανότητα διαχωρισμού των δύο ομάδων, ώστε η πρόβλεψη ταξινόμησης ενός νέου αγωγού να γίνεται με την μεγαλύτερη δυνατή ακρίβεια. Στον ακόλουθο πίνακα, παρουσιάζονται όλα τα σενάρια συνολικά, και τα χαρακτηριστικά που περιλαμβάνονται σε κάθε περίπτωση.

Σενάριο	Χαρακτηριστικά					
1	Length	Diameter	Thickness	Pressure	Life-time	
	DIM1	DIM2	DIM3	DIM4	DIM5	
2	Length	Diameter	Thickness	Pressure	Life-time	
3	Length	Diameter	Thickness	Pressure	Life-time	DIM1
4	Length	Diameter	Thickness	Pressure	Life-time	DIM2
5	Length	Diameter	Thickness	Pressure	Life-time	DIM3
6	Length	Diameter	Thickness	Pressure	Life-time	DIM4
7	Length	Diameter	Thickness	Pressure	Life-time	DIM5
8	Length	Diameter	Thickness	Pressure	Life-time	
	DIM2	DIM3	DIM4	DIM5		
9	Length	Diameter	Thickness	Pressure	Life-time	
	DIM1	DIM3	DIM4	DIM5		
10	Length	Diameter	Thickness	Pressure	Life-time	
	DIM1	DIM2	DIM4	DIM5		
11	Length	Diameter	Thickness	Pressure	Life-time	
	DIM1	DIM2	DIM3	DIM5		
12	Length	Diameter	Thickness	Pressure	Life-time	
	DIM1	DIM2	DIM3	DIM4		
13	DIM1	DIM2	DIM3	DIM4	DIM5	
14	Length	Thickness	Life-time	DIM2	DIM5	
15	Diameter	Pressure	DIM1	DIM3	DIM4	
16	Length	DIM1	DIM2	DIM3	DIM4	DIM5
17	Diameter	DIM1	DIM2	DIM3	DIM4	DIM5
18	Thickness	DIM1	DIM2	DIM3	DIM4	DIM5
19	Pressure	DIM1	DIM2	DIM3	DIM4	DIM5
20	Life-time	DIM1	DIM2	DIM3	DIM4	DIM5
21	Thickness	Life-time	DIM1	DIM2		
22	Length	Pressure	Life-time	DIM3		

6.2 Σενάριο 1°

Στο πρώτο σενάριο, περιλαμβάνονται όλα τα χαρακτηριστικά, συμπεριλαμβανομένων και των ομαδοποιημένων:

Length	DIM1
Diameter	DIM2
Thickness	DIM3
Pressure	DIM4
Life-time	DIM5

- Στον πίνακα 1.1 παρατίθονται τα στατιστικά δεδομένα που προκύπτουν από τις δύο ομάδες, για κάθε χαρακτηριστικό χωριστά.
- Στον πίνακα 1.2 εμφανίζονται τα αποτελέσματα της διαχωριστικής ανάλυσης. Το χαρακτηριστικό με την μεγαλύτερη βαρύτητα και διαχωριστική ισχύ είναι το DIM2, ενώ αυτό με την μικρότερη είναι η διάμετρος (Diameter). Επίσης, όλα τα χαρακτηριστικά παρουσίασαν ευστάθεια, εκτός των Thickness, DIM1 και DIM5 τα οποία είναι ασταθή.
- Στον πίνακα 1.3 παρουσιάζεται ο πίνακας ταξινόμησης, όπου φαίνεται ότι τα ποσοστά σωστής ταξινόμησης, τόσο τα επιμέρους όσο και το συνολικό, είναι αρκετά υψηλά: $EF_1=95,4\%$

  $EF_2=96,6\%$

  $EF_t=95,8\%$
- Το διάγραμμα ταξινόμησης φαίνεται στον πίνακα 1.5 και δείχνει τον υψηλό βαθμό διαχωρισμού που επιτυγχάνεται μεταξύ των δύο ομάδων, καθώς είναι εμφανείς οι διαφορές των πιθανοτήτων P(S/Z) και P(F/Z) για τιμές του Z διάφορες του 0,5.
- Στον πίνακα 1.4 δίνεται η τιμή του Rao's F = 40,6177 για τον έλεγχο ισότητας των μέσων τιμών. Από τους πίνακες της κατανομής F προκύπτει ότι για όλα τα επίπεδα σημαντικότητας, οι μέσες τιμές της μεταβλητής Z των δύο ομάδων δεν είναι ίσες. Επίσης η τιμή του Wilk's Λ = 0,2788 είναι αρκετά χαμηλή, γεγονός που υποδηλώνει την ικανοποιητική διαφορά ανάμεσα στις

μέσες τιμές των δύο ομάδων, και επομένως τον καλό διαχωρισμό που επιτυγχάνεται.

- Στον πίνακα 1.6 δίνεται η τιμή των Box's B = 1149,36 και Box's F = 19,388 και γίνεται ο έλεγχος ισότητας των διασπορών.
- Από τον πίνακα 1.7 προκύπτει ότι ο Κανονικός Συντελεστής Συσχέτισης είναι C.C.C. = 0,8492, τιμή που προσεγγίζει την μονάδα και δηλώνει ότι υπάρχει υψηλή συσχέτιση ανάμεσα στην διαχωριστική συνάρτηση και τις ομάδες. Η αντίστοιχη ιδιοτιμή ισούται με λ = 2,58712.
- Ο έλεγχος προσαρμογής κανονικής κατανομής της μεταβλητής Z, και για τις δύο ομάδες, φαίνεται στον πίνακα 1.8. Προκύπτει ότι σε όλα τα επίπεδα σημαντικότητας (α=1%, 5%, 10%) η μεταβλητή Z ακολουθεί την κανονική κατανομή.
- Τέλος, στον πίνακα 1.9 γίνεται ο έλεγχος παλινδρόμησης. Έτσι, οι τιμές των συντελεστών της εξίσωσης παλινδρόμησης είναι α = -1,6057, b = 3,2637 και R = 0,9915. Ο έλεγχος t-Student είναι θετικός και δείχνει ότι η χρήση του διαγράμματος ταξινόμησης ενδείκνυται για όλο τον πληθυσμό.

6.3 Σενάριο 2°

Στο σενάριο αυτό, περιλαμβάνονται μόνο απλά χαρακτηριστικά, χωρίς να περιλαμβάνονται τα ομαδοποιημένα:

Length
Diameter
Thickness
Pressure
Life-time

- Στον πίνακα 2.1 παρατίθονται τα στατιστικά δεδομένα που προκύπτουν από τις δύο ομάδες, για κάθε χαρακτηριστικό χωριστά.

- Στον πίνακα 2.2 εμφανίζονται τα αποτελέσματα της διαχωριστικής ανάλυσης. Το χαρακτηριστικό με την μεγαλύτερη βαρύτητα και διαχωριστική ισχύ είναι ο χρόνος ζωής (Life-time), ενώ αυτό με την μικρότερη είναι η πίεση (Pressure). Επίσης, όλα τα χαρακτηριστικά παρουσίασαν ευστάθεια, εκτός του πάχους (Thickness).
- Στον πίνακα 2.3 παρουσιάζεται ο πίνακας ταξινόμησης, όπου φαίνεται ότι τα ποσοστά σωστής ταξινόμησης των επιτυχιών έχουν μειωθεί:

$$EF_1=82,6\%$$

$$EF_2=96,6\%$$

$$EF_t=87,5\%$$

- Το διάγραμμα ταξινόμησης φαίνεται στον πίνακα 2.5 και δείχνει τον καλό βαθμό διαχωρισμού που επιτυγχάνεται μεταξύ των δύο ομάδων.
- Στον πίνακα 2.4 δίνεται η τιμή του Rao's F = 34,9822 για τον έλεγχο ισότητας των μέσων τιμών. Από τους πίνακες της κατανομής F προκύπτει ότι για όλα τα επίπεδα σημαντικότητας, οι μέσες τιμές της μεταβλητής Z των δύο ομάδων δεν είναι ίσες. Επίσης η τιμή του Wilk's Λ = 0,4808 δεν είναι αρκετά χαμηλή, γεγονός που υποδηλώνει τον μέτριο διαχωρισμό που επιτυγχάνεται μεταξύ των ομάδων.
- Στον πίνακα 2.6 δίνεται η τιμή των Box's B = 209,63 και Box's F = 13,4636 και γίνεται ο έλεγχος ισότητας των διασπορών.
- Από τον πίνακα 2.7 προκύπτει ότι ο Κανονικός Συντελεστής Συσχέτισης είναι C.C.C. = 0,7205, τιμή όχι ιδιαίτερα ικανοποιητική. Η αντίστοιχη ιδιοτιμή ισούται με λ = 1,0797.
- Ο έλεγχος προσαρμογής κανονικής κατανομής της μεταβλητής Z, και για τις δύο ομάδες, φαίνεται στον πίνακα 2.8. Προκύπτει ότι σε όλα τα επίπεδα σημαντικότητας (α=1%, 5%, 10%) η μεταβλητή Z ακολουθεί την κανονική κατανομή.
- Τέλος, στον πίνακα 2.9 γίνεται ο έλεγχος παλινδρόμησης. Έτσι, οι τιμές των συντελεστών της εξίσωσης παλινδρόμησης είναι α = -0,6967, b = -2,1638 και R = -0,99. Ο έλεγχος t-Student δείχνει ότι η χρήση του διαγράμματος ταξινόμησης δεν ενδείκνυται για όλο τον πληθυσμό.

6.4 Σενάριο 3°

Στο σενάριο αυτό, περιλαμβάνονται όλα τα απλά χαρακτηριστικά, και έχει προστεθεί η ομαδοποιημένη μεταβλητή DIM1:

Length	Pressure
Diameter	Life-time
Thickness	DIM1

- Στον πίνακα 3.1 παρατίθονται τα στατιστικά δεδομένα που προκύπτουν από τις δύο ομάδες, για κάθε χαρακτηριστικό χωριστά.
- Στον πίνακα 3.2 εμφανίζονται τα αποτελέσματα της διαχωριστικής ανάλυσης. Το χαρακτηριστικό με την μεγαλύτερη βαρύτητα και διαχωριστική ισχύ είναι ο χρόνος ζωής (Life-time), ενώ αυτό με την μικρότερη είναι το πάχος (Thickness). Επίσης, όλα τα χαρακτηριστικά παρουσίασαν ευστάθεια, εκτός της πίεσης (Pressure).
- Στον πίνακα 3.3 παρουσιάζεται ο πίνακας ταξινόμησης, όπου φαίνεται ότι τα ποσοστά σωστής ταξινόμησης έχουν μειωθεί:

$$EF_1=82,6\%$$
$$EF_2=94,9\%$$
$$EF_t=86,9\%$$

- Το διάγραμμα ταξινόμησης φαίνεται στον πίνακα 3.5 και δείχνει τον μέτριο βαθμό διαχωρισμού που επιτυγχάνεται μεταξύ των δύο ομάδων.
- Στον πίνακα 3.4 δίνεται η τιμή του Rao's F = 30,321 για τον έλεγχο ισότητας των μέσων τιμών. Από τους πίνακες της κατανομής F προκύπτει ότι για όλα τα επίπεδα σημαντικότητας, οι μέσες τιμές της μεταβλητής Z των δύο ομάδων δεν είναι ίσες. Επίσης η τιμή του Wilk's Λ = 0,4695 δεν είναι αρκετά χαμηλή, γεγονός που υποδηλώνει τον μέτριο διαχωρισμό που επιτυγχάνεται μεταξύ των ομάδων.
- Στον πίνακα 3.6 δίνεται η τιμή των Box's B = 315,317 και Box's F = 14,357 και γίνεται ο έλεγχος ισότητας των διασπορών.

- Από τον πίνακα 3.7 προκύπτει ότι ο Κανονικός Συντελεστής Συσχέτισης είναι C.C.C. = 0,7284, τιμή όχι ιδιαίτερα ικανοποιητική. Η αντίστοιχη ιδιοτιμή ισούται με λ = 1,12997.
- Ο έλεγχος προσαρμογής κανονικής κατανομής της μεταβλητής Z, και για τις δύο ομάδες, φαίνεται στον πίνακα 3.8. Προκύπτει ότι σε όλα τα επίπεδα σημαντικότητας (α=1%, 5%, 10%) η μεταβλητή Z ακολουθεί την κανονική κατανομή.
- Τέλος, στον πίνακα 3.9 γίνεται ο έλεγχος παλινδρόμησης. Έτσι, οι τιμές των συντελεστών της εξίσωσης παλινδρόμησης είναι α = -0,7292, b = 2,2136 και R = 0,99. Ο έλεγχος t-Student δείχνει ότι η χρήση του διαγράμματος ταξινόμησης ενδείκνυται για όλο τον πληθυσμό.

6.5 Σενάριο 4°

Στο σενάριο αυτό, περιλαμβάνονται όλα τα απλά χαρακτηριστικά, και έχει προστεθεί η ομαδοποιημένη μεταβλητή DIM2:

Length	Pressure
Diameter	Life-time
Thickness	DIM2

- Στον πίνακα 4.1 παρατίθονται τα στατιστικά δεδομένα που προκύπτουν από τις δύο ομάδες, για κάθε χαρακτηριστικό χωριστά.
- Στον πίνακα 4.2 εμφανίζονται τα αποτελέσματα της διαχωριστικής ανάλυσης. Το χαρακτηριστικό με την μεγαλύτερη βαρύτητα και διαχωριστική ισχύ είναι το DIM2, ενώ αυτό με την μικρότερη είναι η διάμετρος. Επίσης, όλα τα χαρακτηριστικά παρουσίασαν ευστάθεια, εκτός του πάχους.
- Στον πίνακα 4.3 παρουσιάζεται ο πίνακας ταξινόμησης, όπου φαίνεται ότι τα ποσοστά σωστής ταξινόμησης έχουν αυξηθεί:

$$EF_1=95,4\%$$

$$EF_2=94,9\%$$

EF_t=95,2%

- Το διάγραμμα ταξινόμησης φαίνεται στον πίνακα 4.5 και δείχνει τον υψηλό βαθμό διαχωρισμού που επιτυγχάνεται μεταξύ των δύο ομάδων.
- Στον πίνακα 4.4 δίνεται η τιμή του Rao's F = 55,6321 για τον έλεγχο ισότητας των μέσων τιμών. Από τους πίνακες της κατανομής F προκύπτει ότι για όλα τα επίπεδα σημαντικότητας, οι μέσες τιμές της μεταβλητής Z των δύο ομάδων δεν είναι ίσες. Επίσης η τιμή του Wilk's Λ = 0,3254 είναι σχετικά χαμηλή, γεγονός που υποδηλώνει τον καλό διαχωρισμό που επιτυγχάνεται μεταξύ των ομάδων.
- Στον πίνακα 4.6 δίνεται η τιμή των Box's B = 668,014 και Box's F = 30,417 και γίνεται ο έλεγχος ισότητας των διασπορών.
- Από τον πίνακα 4.7 προκύπτει ότι ο Κανονικός Συντελεστής Συσχέτισης είναι C.C.C. = 0,8213, τιμή η οποία κρίνεται ικανοποιητική. Η αντίστοιχη ιδιοτιμή ισούται με λ = 2,07325.
- Ο έλεγχος προσαρμογής κανονικής κατανομής της μεταβλητής Z, και για τις δύο ομάδες, φαίνεται στον πίνακα 4.8. Προκύπτει ότι σε όλα τα επίπεδα σημαντικότητας (α=1%, 5%, 10%) η μεταβλητή Z ακολουθεί την κανονική κατανομή.
- Τέλος, στον πίνακα 4.9 γίνεται ο έλεγχος παλινδρόμησης. Έτσι, οι τιμές των συντελεστών της εξίσωσης παλινδρόμησης είναι α = -1,3379, b = -2,9984 και R = -0,99. Ο έλεγχος t-Student δείχνει ότι η χρήση του διαγράμματος ταξινόμησης δεν ενδείκνυται για όλο τον πληθυσμό.

6.6 Σενάριο 5°

Στο σενάριο αυτό, περιλαμβάνονται όλα τα απλά χαρακτηριστικά, και έχει προστεθεί η ομαδοποιημένη μεταβλητή DIM3:

Length	Pressure
Diameter	Life-time
Thickness	DIM3

- Στον πίνακα 5.1 παρατίθονται τα στατιστικά δεδομένα που προκύπτουν από τις δύο ομάδες, για κάθε χαρακτηριστικό χωριστά.
- Στον πίνακα 5.2 εμφανίζονται τα αποτελέσματα της διαχωριστικής ανάλυσης. Το χαρακτηριστικό με την μεγαλύτερη βαρύτητα και διαχωριστική ισχύ είναι ο χρόνος ζωής (Life-time), ενώ αυτό με την μικρότερη είναι η διάμετρος (Diameter). Επίσης, όλα τα χαρακτηριστικά παρουσίασαν ευστάθεια, εκτός του πάχους (Thickness).
- Στον πίνακα 5.3 παρουσιάζεται ο πίνακας ταξινόμησης, όπου φαίνεται ότι τα ποσοστά σωστής ταξινόμησης των επιτυχιών έχουν μειωθεί:

  $EF_1=86,2\%$

  $EF_2=94,9\%$

  $EF_t=89,3\%$
- Το διάγραμμα ταξινόμησης φαίνεται στον πίνακα 5.5 και δείχνει τον καλό βαθμό διαχωρισμού που επιτυγχάνεται μεταξύ των δύο ομάδων.
- Στον πίνακα 5.4 δίνεται η τιμή του Rao's F = 34,5247 για τον έλεγχο ισότητας των μέσων τιμών. Από τους πίνακες της κατανομής F προκύπτει ότι για όλα τα επίπεδα σημαντικότητας, οι μέσες τιμές της μεταβλητής Z των δύο ομάδων δεν είναι ίσες. Επίσης η τιμή του Wilk's Λ = 0,4373 δεν είναι αρκετά χαμηλή, γεγονός που υποδηλώνει τον μέτριο διαχωρισμό που επιτυγχάνεται μεταξύ των ομάδων.
- Στον πίνακα 5.6 δίνεται η τιμή των Box's B = 375,267 και Box's F = 17,087 και γίνεται ο έλεγχος ισότητας των διασπορών.

- Από τον πίνακα 5.7 προκύπτει ότι ο Κανονικός Συντελεστής Συσχέτισης είναι C.C.C. = 0,7501, τιμή που θεωρείται ικανοποιητική. Η αντίστοιχη ιδιοτιμή ισούται με λ = 1,28663.
- Ο έλεγχος προσαρμογής κανονικής κατανομής της μεταβλητής Z, και για τις δύο ομάδες, φαίνεται στον πίνακα 5.8. Προκύπτει ότι σε όλα τα επίπεδα σημαντικότητας (α=1%, 5%, 10%) η μεταβλητή Z ακολουθεί την κανονική κατανομή.
- Τέλος, στον πίνακα 5.9 γίνεται ο έλεγχος παλινδρόμησης. Έτσι, οι τιμές των συντελεστών της εξίσωσης παλινδρόμησης είναι α = -0,8303, b = -2,3621 και R = -0,989. Ο έλεγχος t-Student δείχνει ότι η χρήση του διαγράμματος ταξινόμησης δεν ενδείκνυται για όλο τον πληθυσμό.

6.7 Σενάριο 6°

Στο σενάριο αυτό, περιλαμβάνονται όλα τα απλά χαρακτηριστικά, και έχει προστεθεί η ομαδοποιημένη μεταβλητή DIM4:

Length	Pressure
Diameter	Life-time
Thickness	DIM4

- Στον πίνακα 6.1 παρατίθονται τα στατιστικά δεδομένα που προκύπτουν από τις δύο ομάδες, για κάθε χαρακτηριστικό χωριστά.
- Στον πίνακα 6.2 εμφανίζονται τα αποτελέσματα της διαχωριστικής ανάλυσης. Το χαρακτηριστικό με την μεγαλύτερη βαρύτητα και διαχωριστική ισχύ είναι ο χρόνος ζωής (Life-time), ενώ αυτό με την μικρότερη είναι το πάχος (Thickness). Επίσης, όλα τα χαρακτηριστικά παρουσίασαν ευστάθεια, εκτός της πίεσης (Pressure) και του πάχους (Thickness).
- Στον πίνακα 6.3 παρουσιάζεται ο πίνακας ταξινόμησης, όπου φαίνεται ότι τα ποσοστά σωστής ταξινόμησης έχουν αυξηθεί:

EF_1=89,0%

EF_2=98,3%

EF_t=92,3%

- Το διάγραμμα ταξινόμησης φαίνεται στον πίνακα 6.5 και δείχνει τον καλό διαχωρισμό που επιτυγχάνεται μεταξύ των δύο ομάδων.
- Στον πίνακα 6.4 δίνεται η τιμή του Rao's F = 36,2315 για τον έλεγχο ισότητας των μέσων τιμών. Από τους πίνακες της κατανομής F προκύπτει ότι για όλα τα επίπεδα σημαντικότητας, οι μέσες τιμές της μεταβλητής Z των δύο ομάδων δεν είναι ίσες. Επίσης η τιμή του Wilk's Λ = 0,4255 υποδηλώνει τον μέτριο διαχωρισμό που επιτυγχάνεται μεταξύ των ομάδων.
- Στον πίνακα 6.6 δίνεται η τιμή των Box's B = 261,663 και Box's F = 11,914 και γίνεται ο έλεγχος ισότητας των διασπορών.
- Από τον πίνακα 6.7 προκύπτει ότι ο Κανονικός Συντελεστής Συσχέτισης είναι C.C.C. = 0,758. Η αντίστοιχη ιδιοτιμή ισούται με λ = 1,35024.
- Ο έλεγχος προσαρμογής κανονικής κατανομής της μεταβλητής Z, και για τις δύο ομάδες, φαίνεται στον πίνακα 6.8. Προκύπτει ότι σε όλα τα επίπεδα σημαντικότητας (α=1%, 5%, 10%) η μεταβλητή Z ακολουθεί την κανονική κατανομή.
- Τέλος, στον πίνακα 6.9 γίνεται ο έλεγχος παλινδρόμησης. Έτσι, οι τιμές των συντελεστών της εξίσωσης παλινδρόμησης είναι α = -0,8713, b = 2,4198 και R = 0,999. Ο έλεγχος t-Student δείχνει ότι η χρήση του διαγράμματος ταξινόμησης ενδείκνυται για όλο τον πληθυσμό.

6.8 Σενάριο 7°

Στο σενάριο αυτό, περιλαμβάνονται όλα τα απλά χαρακτηριστικά, και έχει προστεθεί η ομαδοποιημένη μεταβλητή DIM5:

Length	Pressure
Diameter	Life-time
Thickness	DIM5

- Στον πίνακα 7.1 παρατίθονται τα στατιστικά δεδομένα που προκύπτουν από τις δύο ομάδες, για κάθε χαρακτηριστικό χωριστά.
- Στον πίνακα 7.2 εμφανίζονται τα αποτελέσματα της διαχωριστικής ανάλυσης. Το χαρακτηριστικό με την μεγαλύτερη βαρύτητα και διαχωριστική ισχύ είναι ο χρόνος ζωής (Life-time), ενώ αυτό με την μικρότερη είναι το DIM5. Επίσης, όλα τα χαρακτηριστικά παρουσίασαν ευστάθεια, εκτός του πάχους και της μεταβλητής DIM5.
- Στον πίνακα 7.3 παρουσιάζεται ο πίνακας ταξινόμησης, όπου φαίνεται ότι τα ποσοστά σωστής ταξινόμησης έχουν μειωθεί:

$$EF_1=82,6\%$$
$$EF_2=94,9\%$$
$$EF_t=87,5\%$$

- Το διάγραμμα ταξινόμησης φαίνεται στον πίνακα 7.5 και δείχνει τον μέτριο βαθμό διαχωρισμού που επιτυγχάνεται μεταξύ των δύο ομάδων.
- Στον πίνακα 7.4 δίνεται η τιμή του Rao's F = 29,0479 για τον έλεγχο ισότητας των μέσων τιμών. Από τους πίνακες της κατανομής F προκύπτει ότι για όλα τα επίπεδα σημαντικότητας, οι μέσες τιμές της μεταβλητής Z των δύο ομάδων δεν είναι ίσες. Επίσης η τιμή του Wilk's Λ = 0,4802 δεν είναι αρκετά χαμηλή, γεγονός που υποδηλώνει τον μέτριο διαχωρισμό που επιτυγχάνεται μεταξύ των ομάδων.
- Στον πίνακα 7.8 δίνεται η τιμή των Box's B = 378,218 και Box's F = 17,221 και γίνεται ο έλεγχος ισότητας των διασπορών.
- Από τον πίνακα 7.9 προκύπτει ότι ο Κανονικός Συντελεστής Συσχέτισης είναι C.C.C. = 0,7201. Η αντίστοιχη ιδιοτιμή ισούται με λ = 1,08253.

- Ο έλεγχος προσαρμογής κανονικής κατανομής της μεταβλητής Z, και για τις δύο ομάδες, φαίνεται στον πίνακα 7.6. Προκύπτει ότι σε όλα τα επίπεδα σημαντικότητας (α=1%, 5%, 10%) η μεταβλητή Z ακολουθεί την κανονική κατανομή.
- Τέλος, στον πίνακα 7.7 γίνεται ο έλεγχος παλινδρόμησης. Έτσι, οι τιμές των συντελεστών της εξίσωσης παλινδρόμησης είναι α = -0,6986, b = 2,1667 και R = 0,999. Ο έλεγχος t-Student δείχνει ότι η χρήση του διαγράμματος ταξινόμησης ενδείκνυται για όλο τον πληθυσμό.

6.9 Σενάριο 8°

Στο σενάριο αυτό, περιλαμβάνονται όλα τα χαρακτηριστικά, και έχει παραλειφθεί η ομαδοποιημένη μεταβλητή DIM1:

Length	DIM2
Diameter	DIM3
Thickness	DIM4
Pressure	DIM5
Life-time	

- Στον πίνακα 8.1 παρατίθονται τα στατιστικά δεδομένα που προκύπτουν από τις δύο ομάδες, για κάθε χαρακτηριστικό χωριστά.
- Στον πίνακα 8.2 εμφανίζονται τα αποτελέσματα της διαχωριστικής ανάλυσης. Το χαρακτηριστικό με την μεγαλύτερη βαρύτητα και διαχωριστική ισχύ είναι το DIM2, ενώ αυτό με την μικρότερη είναι η διάμετρος. Επίσης, τα χαρακτηριστικά διάμετρος, πάχος και DIM5 παρουσίασαν αστάθεια.
- Στον πίνακα 8.3 παρουσιάζεται ο πίνακας ταξινόμησης, όπου φαίνονται τα αρκετά ικανοποιητικά ποσοστά σωστής ταξινόμησης :

EF_1=96,3%

EF_2=96,6%

EF_t=96,4%

- Το διάγραμμα ταξινόμησης φαίνεται στον πίνακα 8.5 και δείχνει τον αρκετά καλό διαχωρισμό που επιτυγχάνεται μεταξύ των δύο ομάδων.
- Στον πίνακα 8.4 δίνεται η τιμή του Rao's F = 43,8844 για τον έλεγχο ισότητας των μέσων τιμών. Από τους πίνακες της κατανομής F προκύπτει ότι για όλα τα επίπεδα σημαντικότητας, οι μέσες τιμές της μεταβλητής Z των δύο ομάδων δεν είναι ίσες. Επίσης η τιμή του Wilk's Λ = 0,2857 υποδηλώνει τον αρκετά καλό διαχωρισμό που επιτυγχάνεται μεταξύ των ομάδων.
- Στον πίνακα 8.7 δίνεται η τιμή των Box's B = 1032,198 και Box's F = 21,444 και γίνεται ο έλεγχος ισότητας των διασπορών.
- Από τον πίνακα 8.6 προκύπτει ότι ο Κανονικός Συντελεστής Συσχέτισης είναι C.C.C. = 0,8451, τιμή ιδιαίτερα ικανοποιητική. Η αντίστοιχη ιδιοτιμή ισούται με λ = 2,49974.
- Ο έλεγχος προσαρμογής κανονικής κατανομής της μεταβλητής Z, και για τις δύο ομάδες, φαίνεται στον πίνακα 8.8. Προκύπτει ότι σε όλα τα επίπεδα σημαντικότητας (α=1%, 5%, 10%) η μεταβλητή Z ακολουθεί την κανονική κατανομή.
- Τέλος, στον πίνακα 8.9 γίνεται ο έλεγχος παλινδρόμησης. Έτσι, οι τιμές των συντελεστών της εξίσωσης παλινδρόμησης είναι α = -1,612, b = 3,2909 και R = 0,999. Ο έλεγχος t-Student δείχνει ότι η χρήση του διαγράμματος ταξινόμησης ενδείκνυται για όλο τον πληθυσμό.

6.10 Σενάριο 9°

Στο σενάριο αυτό, περιλαμβάνονται όλα τα χαρακτηριστικά, και έχει παραλειφθεί η ομαδοποιημένη μεταβλητή DIM2:

Length	DIM1
Diameter	DIM3
Thickness	DIM4
Pressure	DIM5
Life-time	

- Στον πίνακα 9.1 παρατίθονται τα στατιστικά δεδομένα που προκύπτουν από τις δύο ομάδες, για κάθε χαρακτηριστικό χωριστά.
- Στον πίνακα 9.2 εμφανίζονται τα αποτελέσματα της διαχωριστικής ανάλυσης. Το χαρακτηριστικό με την μεγαλύτερη βαρύτητα και διαχωριστική ισχύ είναι ο χρόνος ζωής (Life-time), ενώ αυτό με την μικρότερη είναι η διάμετρος (Diameter). Επίσης, όλα τα χαρακτηριστικά παρουσίασαν ευστάθεια, εκτός του πάχους και του DIM1.
- Στον πίνακα 9.3 παρουσιάζεται ο πίνακας ταξινόμησης, όπου φαίνεται ότι τα ποσοστά σωστής ταξινόμησης έχουν μειωθεί:

$$EF_1=89,0\%$$
$$EF_2=94,9\%$$
$$EF_t=91,1\%$$

- Το διάγραμμα ταξινόμησης φαίνεται στον πίνακα 9.5 και δείχνει τον καλό βαθμό διαχωρισμού που επιτυγχάνεται μεταξύ των δύο ομάδων.
- Στον πίνακα 9.4 δίνεται η τιμή του Rao's F = 25,9365 για τον έλεγχο ισότητας των μέσων τιμών. Από τους πίνακες της κατανομής F προκύπτει ότι για όλα τα επίπεδα σημαντικότητας, οι μέσες τιμές της μεταβλητής Z των δύο ομάδων δεν είναι ίσες. Επίσης η τιμή του Wilk's Λ = 0,4036 δείχνει τον σχετικά καλό διαχωρισμό που επιτυγχάνεται μεταξύ των ομάδων.
- Στον πίνακα 9.6 δίνεται η τιμή των Box's B = 683,249 και Box's F = 14,195 και γίνεται ο έλεγχος ισότητας των διασπορών.

- Από τον πίνακα 9.7 προκύπτει ότι ο Κανονικός Συντελεστής Συσχέτισης είναι C.C.C. = 0,7722, εκφράζοντας τον ικανοποιητικό βαθμό συσχέτισης μεταξύ της διαχωριστικής συνάρτησης και των ομάδων. Η αντίστοιχη ιδιοτιμή ισούται με λ = 1,4774.
- Ο έλεγχος προσαρμογής κανονικής κατανομής της μεταβλητής Z, και για τις δύο ομάδες, φαίνεται στον πίνακα 9.8. Προκύπτει ότι σε όλα τα επίπεδα σημαντικότητας (α=1%, 5%, 10%) η μεταβλητή Z ακολουθεί την κανονική κατανομή.
- Τέλος, στον πίνακα 9.9 γίνεται ο έλεγχος παλινδρόμησης. Έτσι, οι τιμές των συντελεστών της εξίσωσης παλινδρόμησης είναι α = -0,9534, b = -2,5312 και R = -0,989. Ο έλεγχος t-Student δείχνει ότι η χρήση του διαγράμματος ταξινόμησης δεν ενδείκνυται για όλο τον πληθυσμό.

6.11 Σενάριο 10°

Στο σενάριο αυτό, περιλαμβάνονται όλα τα χαρακτηριστικά, και έχει παραλειφθεί η ομαδοποιημένη μεταβλητή DIM3:

Length	DIM1
Diameter	DIM2
Thickness	DIM4
Pressure	DIM5
Life-time	

- Στον πίνακα 10.1 παρατίθονται τα στατιστικά δεδομένα που προκύπτουν από τις δύο ομάδες, για κάθε χαρακτηριστικό χωριστά.
- Στον πίνακα 10.2 εμφανίζονται τα αποτελέσματα της διαχωριστικής ανάλυσης. Το χαρακτηριστικό με την μεγαλύτερη βαρύτητα και διαχωριστική ισχύ είναι το DIM2, ενώ αυτό με την μικρότερη είναι η διάμετρος. Επίσης, τα χαρακτηριστικά πάχος, DIM1 και DIM5 παρουσίασαν αστάθεια.

- Στον πίνακα 10.3 παρουσιάζεται ο πίνακας ταξινόμησης, όπου φαίνονται τα αρκετά ικανοποιητικά ποσοστά σωστής ταξινόμησης :

$EF_1=95,4\%$

$EF_2=98,3\%$

$EF_t=96,4\%$

- Το διάγραμμα ταξινόμησης φαίνεται στον πίνακα 10.5 και δείχνει τον αρκετά καλό διαχωρισμό που επιτυγχάνεται μεταξύ των δύο ομάδων.
- Στον πίνακα 10.4 δίνεται η τιμή του Rao's F = 45,2233 για τον έλεγχο ισότητας των μέσων τιμών. Από τους πίνακες της κατανομής F προκύπτει ότι για όλα τα επίπεδα σημαντικότητας, οι μέσες τιμές της μεταβλητής Z των δύο ομάδων δεν είναι ίσες. Επίσης η τιμή του Wilk's Λ = 0,2796 υποδηλώνει τον αρκετά καλό διαχωρισμό που επιτυγχάνεται μεταξύ των ομάδων.
- Στον πίνακα 10.6 δίνεται η τιμή των Box's B = 998,169 και Box's F = 20,737 και γίνεται ο έλεγχος ισότητας των διασπορών.
- Από τον πίνακα 10.7 προκύπτει ότι ο Κανονικός Συντελεστής Συσχέτισης είναι C.C.C. = 0,8487, τιμή ιδιαίτερα ικανοποιητική. Η αντίστοιχη ιδιοτιμή ισούται με λ = 2,57601.
- Ο έλεγχος προσαρμογής κανονικής κατανομής της μεταβλητής Z, και για τις δύο ομάδες, φαίνεται στον πίνακα 10.8. Προκύπτει ότι σε όλα τα επίπεδα σημαντικότητας (α=1%, 5%, 10%) η μεταβλητή Z ακολουθεί την κανονική κατανομή.
- Τέλος, στον πίνακα 10.9 γίνεται ο έλεγχος παλινδρόμησης. Έτσι, οι τιμές των συντελεστών της εξίσωσης παλινδρόμησης είναι α = -1,5995, b = 3,2595 και R = 0,992. Ο έλεγχος t-Student δείχνει ότι η χρήση του διαγράμματος ταξινόμησης ενδείκνυται για όλο τον πληθυσμό.

6.12 Σενάριο 11°

Στο σενάριο αυτό, περιλαμβάνονται όλα τα χαρακτηριστικά, και έχει παραλειφθεί η ομαδοποιημένη μεταβλητή DIM4:

Length	DIM1
Diameter	DIM2
Thickness	DIM3
Pressure	DIM5
Life-time	

- Στον πίνακα 11.1 παρατίθονται τα στατιστικά δεδομένα που προκύπτουν από τις δύο ομάδες, για κάθε χαρακτηριστικό χωριστά.
- Στον πίνακα 11.2 εμφανίζονται τα αποτελέσματα της διαχωριστικής ανάλυσης. Το χαρακτηριστικό με την μεγαλύτερη βαρύτητα και διαχωριστική ισχύ είναι το DIM2, ενώ αυτό με την μικρότερη είναι η διάμετρος. Επίσης, τα χαρακτηριστικά πάχος DIM1 και DIM5 παρουσίασαν αστάθεια.
- Στον πίνακα 11.3 παρουσιάζεται ο πίνακας ταξινόμησης, όπου φαίνεται η μείωση του ποσοστού σωστής ταξινόμησης των αποτυχιών:

$$EF_1=96,3\%$$
$$EF_2=94,9\%$$
$$EF_t=95,8\%$$

- Το διάγραμμα ταξινόμησης φαίνεται στον πίνακα 11.5 και δείχνει τον αρκετά καλό διαχωρισμό που επιτυγχάνεται μεταξύ των δύο ομάδων.
- Στον πίνακα 11.4 δίνεται η τιμή του Rao's F = 45,146 για τον έλεγχο ισότητας των μέσων τιμών. Από τους πίνακες της κατανομής F προκύπτει ότι για όλα τα επίπεδα σημαντικότητας, οι μέσες τιμές της μεταβλητής Z των δύο ομάδων δεν είναι ίσες. Επίσης η τιμή του Wilk's Λ = 0,280 υποδηλώνει τον αρκετά καλό διαχωρισμό που επιτυγχάνεται μεταξύ των ομάδων.
- Στον πίνακα 11.6 δίνεται η τιμή των Box's B = 1120,326 και Box's F = 23,275 και γίνεται ο έλεγχος ισότητας των διασπορών.

- Από τον πίνακα 11.7 προκύπτει ότι ο Κανονικός Συντελεστής Συσχέτισης είναι C.C.C. = 0,8485, τιμή ιδιαίτερα ικανοποιητική. Η αντίστοιχη ιδιοτιμή ισούται με λ = 2,57161.
- Ο έλεγχος προσαρμογής κανονικής κατανομής της μεταβλητής Z, και για τις δύο ομάδες, φαίνεται στον πίνακα 11.8. Προκύπτει ότι σε όλα τα επίπεδα σημαντικότητας (α=1%, 5%, 10%) η μεταβλητή Z ακολουθεί την κανονική κατανομή.
- Τέλος, στον πίνακα 11.9 γίνεται ο έλεγχος παλινδρόμησης. Έτσι, οι τιμές των συντελεστών της εξίσωσης παλινδρόμησης είναι α = -1,5961, b = 3,2531 και R = 0,9913. Ο έλεγχος t-Student δείχνει ότι η χρήση του διαγράμματος ταξινόμησης ενδείκνυται για όλο τον πληθυσμό.

6.13 Σενάριο 12°

Στο σενάριο αυτό, περιλαμβάνονται όλα τα χαρακτηριστικά, και έχει παραλειφθεί η ομαδοποιημένη μεταβλητή DIM5:

Length	DIM1
Diameter	DIM2
Thickness	DIM3
Pressure	DIM4
Life-time	

- Στον πίνακα 12.1 παρατίθονται τα στατιστικά δεδομένα που προκύπτουν από τις δύο ομάδες, για κάθε χαρακτηριστικό χωριστά.
- Στον πίνακα 12.2 εμφανίζονται τα αποτελέσματα της διαχωριστικής ανάλυσης. Το χαρακτηριστικό με την μεγαλύτερη βαρύτητα και διαχωριστική ισχύ είναι το DIM2, ενώ αυτό με την μικρότερη είναι η διάμετρος. Επίσης, τα χαρακτηριστικά πάχος και DIM1 παρουσίασαν αστάθεια.
- Στον πίνακα 12.3 παρουσιάζεται ο πίνακας ταξινόμησης, όπου φαίνονται τα ακόλουθα ποσοστά σωστής ταξινόμησης :

$$EF_1=94,5\%$$

EF_2=96,6%

EF_t=95,2%

- Το διάγραμμα ταξινόμησης φαίνεται στον πίνακα 12.5 και δείχνει τον αρκετά καλό διαχωρισμό που επιτυγχάνεται μεταξύ των δύο ομάδων.
- Στον πίνακα 12.4 δίνεται η τιμή του Rao's F = 40,933 για τον έλεγχο ισότητας των μέσων τιμών. Από τους πίνακες της κατανομής F προκύπτει ότι για όλα τα επίπεδα σημαντικότητας, οι μέσες τιμές της μεταβλητής Z των δύο ομάδων δεν είναι ίσες. Επίσης η τιμή του Wilk's Λ = 0,3002 υποδηλώνει τον καλό διαχωρισμό που επιτυγχάνεται μεταξύ των ομάδων.
- Στον πίνακα 12.6 δίνεται η τιμή των Box's B = 999,7413 και Box's F = 20,7705 και γίνεται ο έλεγχος ισότητας των διασπορών.
- Από τον πίνακα 12.7 προκύπτει ότι ο Κανονικός Συντελεστής Συσχέτισης είναι C.C.C. = 0,8366, τιμή ιδιαίτερα ικανοποιητική. Η αντίστοιχη ιδιοτιμή ισούται με λ = 2,33163.
- Ο έλεγχος προσαρμογής κανονικής κατανομής της μεταβλητής Z, και για τις δύο ομάδες, φαίνεται στον πίνακα 12.8. Προκύπτει ότι σε όλα τα επίπεδα σημαντικότητας (α=1%, 5%, 10%) η μεταβλητή Z ακολουθεί την κανονική κατανομή.
- Τέλος, στον πίνακα 12.9 γίνεται ο έλεγχος παλινδρόμησης. Έτσι, οι τιμές των συντελεστών της εξίσωσης παλινδρόμησης είναι α = -1,5054, b = 3,1807 και R = 0,999. Ο έλεγχος t-Student δείχνει ότι η χρήση του διαγράμματος ταξινόμησης ενδείκνυται για όλο τον πληθυσμό.

6.14 Σενάριο 13°

Στο σενάριο αυτό, περιλαμβάνονται όλα τα ομαδοποιημένα χαρακτηριστικά:

DIM1
DIM2
DIM3
DIM4
DIM5

- Στον πίνακα 13.1 παρατίθονται τα στατιστικά δεδομένα που προκύπτουν από τις δύο ομάδες, για κάθε χαρακτηριστικό χωριστά.
- Στην συνέχεια, η μαθηματική διαδικασία διακόπτεται γιατί δεν υπάρχει ο αντίστροφος πίνακας του W - Within Groups SSCP. Επομένως είναι αδύνατος ο προσδιορισμός της διαχωριστικής συνάρτησης και των άλλων παραμέτρων της ανάλυσης.

6.15 Σενάριο 14°

Στο σενάριο αυτό, περιλαμβάνονται εκείνα τα χαρακτηριστικά που κατατάσσονται τα 5 πρώτα στο 1° σενάριο, όπως φαίνεται στον πίνακα 1.2:

Length
Thickness
Life-time
DIM2
DIM5

- Στον πίνακα 14.1 παρατίθονται τα στατιστικά δεδομένα που προκύπτουν από τις δύο ομάδες, για κάθε χαρακτηριστικό χωριστά.

- Στον πίνακα 14.2 εμφανίζονται τα αποτελέσματα της διαχωριστικής ανάλυσης. Το χαρακτηριστικό με την μεγαλύτερη βαρύτητα και διαχωριστική ισχύ είναι το DIM2, ενώ αυτό με την μικρότερη είναι το πάχος. Επίσης, τα χαρακτηριστικά πάχος και DIM5 παρουσίασαν αστάθεια.
- Στον πίνακα 14.3 παρουσιάζεται ο πίνακας ταξινόμησης, όπου φαίνονται τα ακόλουθα ποσοστά σωστής ταξινόμησης :

$$EF_1=96,3\%$$
$$EF_2=94,9\%$$
$$EF_t=95,8\%$$

- Το διάγραμμα ταξινόμησης φαίνεται στον πίνακα 14.5 και δείχνει τον αρκετά καλό διαχωρισμό που επιτυγχάνεται μεταξύ των δύο ομάδων.
- Στον πίνακα 14.4 δίνεται η τιμή του Rao's F = 77,9246 για τον έλεγχο ισότητας των μέσων τιμών. Από τους πίνακες της κατανομής F προκύπτει ότι για όλα τα επίπεδα σημαντικότητας, οι μέσες τιμές της μεταβλητής Z των δύο ομάδων δεν είναι ίσες. Επίσης η τιμή του Wilk's Λ = 0,2937 υποδηλώνει τον καλό διαχωρισμό που επιτυγχάνεται μεταξύ των ομάδων.
- Στον πίνακα 12.6 δίνεται η τιμή των Box's B = 650,5723 και Box's F = 41,781 και γίνεται ο έλεγχος ισότητας των διασπορών.
- Από τον πίνακα 14.7 προκύπτει ότι ο Κανονικός Συντελεστής Συσχέτισης είναι C.C.C. = 0,8404. Η αντίστοιχη ιδιοτιμή ισούται με λ = 2,40508.
- Ο έλεγχος προσαρμογής κανονικής κατανομής της μεταβλητής Z, και για τις δύο ομάδες, φαίνεται στον πίνακα 14.8. Προκύπτει ότι σε όλα τα επίπεδα σημαντικότητας (α=1%, 5%, 10%) η μεταβλητή Z ακολουθεί την κανονική κατανομή.
- Τέλος, στον πίνακα 14.9 γίνεται ο έλεγχος παλινδρόμησης. Έτσι, οι τιμές των συντελεστών της εξίσωσης παλινδρόμησης είναι α = -1,5523, b = 3,2299 και R = 0,999. Ο έλεγχος t-Student δείχνει ότι η χρήση του διαγράμματος ταξινόμησης ενδείκνυται για όλο τον πληθυσμό.

6.16 Σενάριο 15°

Στο σενάριο αυτό, περιλαμβάνονται εκείνα τα χαρακτηριστικά που κατατάσσονται τα 5 τελευταία στο 1° σενάριο, όπως φαίνεται στον πίνακα 1.2:

Diameter
Pressure
DIM1
DIM3
DIM4

- Στον πίνακα 15.1 παρατίθονται τα στατιστικά δεδομένα που προκύπτουν από τις δύο ομάδες, για κάθε χαρακτηριστικό χωριστά.
- Στον πίνακα 15.2 εμφανίζονται τα αποτελέσματα της διαχωριστικής ανάλυσης. Το χαρακτηριστικό με την μεγαλύτερη βαρύτητα και διαχωριστική ισχύ είναι η πίεση (Pressure), ενώ αυτό με την μικρότερη είναι το DIM1. Επίσης, τα χαρακτηριστικά διάμετρος και DIM1 παρουσίασαν αστάθεια.
- Στον πίνακα 15.3 παρουσιάζεται ο πίνακας ταξινόμησης, όπου φαίνονται τα ακόλουθα ποσοστά σωστής ταξινόμησης :

$EF_1=65,1\%$

$EF_2=84,7\%$

$EF_t=72,0\%$

- Το διάγραμμα ταξινόμησης φαίνεται στον πίνακα 15.5 και δείχνει ότι ο διαχωρισμός που επιτυγχάνεται μεταξύ των δύο ομάδων είναι πολύ ασθενής.
- Στον πίνακα 15.4 δίνεται η τιμή του Rao's F = 5,4478, η οποία όπως φαίνεται είναι και πολύ μικρή σε σχέση με τα άλλα σενάρια. Από τους πίνακες της κατανομής F προκύπτει ότι για όλα τα επίπεδα σημαντικότητας, οι μέσες τιμές της μεταβλητής Z των δύο ομάδων δεν είναι ίσες. Επίσης η τιμή του Wilk's Λ = 0,8561 υποδηλώνει τον πολύ μικρό διαχωρισμό που επιτυγχάνεται μεταξύ των ομάδων.
- Στον πίνακα 15.6 δίνεται η τιμή των Box's B = 234,487 και Box's F = 15,059 και γίνεται ο έλεγχος ισότητας των διασπορών.

- Από τον πίνακα 15.7 προκύπτει ότι ο Κανονικός Συντελεστής Συσχέτισης είναι C.C.C. = 0,3794. Η τιμή αυτή του συντελεστή επιβεβαιώνει την υποψία της ασθενούς συσχέτισης μεταξύ της διαχωριστικής συνάρτησης και των δύο ομάδων. Η αντίστοιχη ιδιοτιμή ισούται με λ = 0,16814.
- Ο έλεγχος προσαρμογής κανονικής κατανομής της μεταβλητής Z, και για τις δύο ομάδες, φαίνεται στον πίνακα 15.8. Προκύπτει ότι σε όλα τα επίπεδα σημαντικότητας (α=1%, 5%, 10%) η μεταβλητή Z ακολουθεί την κανονική κατανομή.
- Τέλος, στον πίνακα 15.9 γίνεται ο έλεγχος παλινδρόμησης. Έτσι, οι τιμές των συντελεστών της εξίσωσης παλινδρόμησης είναι α = -0,1085, b = 0,8539 και R = 0,999. Ο έλεγχος t-Student δείχνει ότι η χρήση του διαγράμματος ταξινόμησης ενδείκνυται για όλο τον πληθυσμό.

6.17 Σενάριο 16°

Στο σενάριο αυτό, περιλαμβάνονται όλα τα ομαδοποιημένα χαρακτηριστικά, έχοντας προσθέσει την μεταβλητή Length:

Length
DIM1
DIM2
DIM3
DIM4
DIM5

- Στον πίνακα 16.1 παρατίθονται τα στατιστικά δεδομένα που προκύπτουν από τις δύο ομάδες, για κάθε χαρακτηριστικό χωριστά.
- Στον πίνακα 16.2 εμφανίζονται τα αποτελέσματα της διαχωριστικής ανάλυσης. Το χαρακτηριστικό με την μεγαλύτερη βαρύτητα και διαχωριστική ισχύ είναι το DIM2, ενώ αυτό με την μικρότερη είναι το DIM3. Όλα τα χαρακτηριστικά του σεναρίου εμφανίζονται ευσταθή.

- Στον πίνακα 16.3 παρουσιάζεται ο πίνακας ταξινόμησης, όπου φαίνονται τα ακόλουθα ποσοστά σωστής ταξινόμησης :

$$EF_1=90,8\%$$
$$EF_2=86,4\%$$
$$EF_t=89,3\%$$

- Το διάγραμμα ταξινόμησης φαίνεται στον πίνακα 16.5 και δείχνει ότι ο διαχωρισμός που επιτυγχάνεται μεταξύ των δύο ομάδων δεν είναι ιδιαίτερα ικανοποιητικός.
- Στον πίνακα 16.4 δίνεται η τιμή του Rao's F = 23,8655 για τον έλεγχο ισότητας των μέσων τιμών. Από τους πίνακες της κατανομής F προκύπτει ότι για όλα τα επίπεδα σημαντικότητας, οι μέσες τιμές της μεταβλητής Z των δύο ομάδων δεν είναι ίσες. Επίσης η τιμή του Wilk's Λ = 0,5293 υποδηλώνει τον μικρό διαχωρισμό που επιτυγχάνει η διαχωριστική συνάρτηση.
- Στον πίνακα 16.6 δίνεται η τιμή των Box's B = 694,063 και Box's F = 31,603 και γίνεται ο έλεγχος ισότητας των διασπορών.
- Από τον πίνακα 16.7 προκύπτει ότι ο Κανονικός Συντελεστής Συσχέτισης είναι C.C.C. = 0,6861, ενισχύοντας το συμπέρασμα που προέκυψε από την τιμή του Wilk's Λ. Η αντίστοιχη ιδιοτιμή ισούται με λ = 0,8894.
- Ο έλεγχος προσαρμογής κανονικής κατανομής της μεταβλητής Z, και για τις δύο ομάδες, φαίνεται στον πίνακα 16.8. Προκύπτει ότι σε όλα τα επίπεδα σημαντικότητας (α=1%, 5%, 10%) η μεταβλητή Z ακολουθεί την κανονική κατανομή.
- Τέλος, στον πίνακα 16.9 γίνεται ο έλεγχος παλινδρόμησης. Έτσι, οι τιμές των συντελεστών της εξίσωσης παλινδρόμησης είναι α = -0,5739, b = -1,9639 και R = -0,99. Ο έλεγχος t-Student δείχνει ότι η χρήση του διαγράμματος ταξινόμησης δεν ενδείκνυται για όλο τον πληθυσμό.

6.18 Σενάριο 17°

Στο σενάριο αυτό, περιλαμβάνονται όλα τα ομαδοποιημένα χαρακτηριστικά, έχοντας προσθέσει την μεταβλητή Diameter:

Diameter
DIM1
DIM2
DIM3
DIM4
DIM5

- Στον πίνακα 17.1 παρατίθονται τα στατιστικά δεδομένα που προκύπτουν από τις δύο ομάδες, για κάθε χαρακτηριστικό χωριστά.
- Στον πίνακα 17.2 εμφανίζονται τα αποτελέσματα της διαχωριστικής ανάλυσης. Το χαρακτηριστικό με την μεγαλύτερη βαρύτητα και διαχωριστική ισχύ είναι το DIM2, ενώ αυτό με την μικρότερη είναι το DIM3. Επίσης, όλα τα χαρακτηριστικά είναι ευσταθή εκτός της διαμέτρου που παρουσίασε αστάθεια.
- Στον πίνακα 17.3 παρουσιάζεται ο πίνακας ταξινόμησης, όπου φαίνονται τα ακόλουθα, χαμηλά, ποσοστά σωστής ταξινόμησης :

$EF_1=76,3\%$

$EF_2=86,4\%$

$EF_t=79,8\%$

- Το διάγραμμα ταξινόμησης φαίνεται στον πίνακα 17.5 και δείχνει τον ασθενή διαχωρισμό που επιτυγχάνεται μεταξύ των δύο ομάδων.
- Στον πίνακα 17.4 δίνεται η τιμή του Rao's F = 13,2087 για τον έλεγχο ισότητας των μέσων τιμών. Από τους πίνακες της κατανομής F προκύπτει ότι για όλα τα επίπεδα σημαντικότητας, οι μέσες τιμές της μεταβλητής Z των δύο ομάδων δεν είναι ίσες. Επίσης η τιμή του Wilk's Λ = 0,6701 επιβεβαιώνει την υποψία ασθενή διαχωρισμού μεταξύ των δύο ομάδων.
- Στον πίνακα 17.6 δίνεται η τιμή των Box's B = 418,454 και Box's F = 19,054 και γίνεται ο έλεγχος ισότητας των διασπορών.

- Από τον πίνακα 17.7 προκύπτει ότι ο Κανονικός Συντελεστής Συσχέτισης είναι C.C.C. = 0,5743. Η αντίστοιχη ιδιοτιμή ισούται με λ = 0,49225.
- Ο έλεγχος προσαρμογής κανονικής κατανομής της μεταβλητής Z, και για τις δύο ομάδες, φαίνεται στον πίνακα 17.8. Προκύπτει ότι σε όλα τα επίπεδα σημαντικότητας (α=1%, 5%, 10%) η μεταβλητή Z ακολουθεί την κανονική κατανομή.
- Τέλος, στον πίνακα 17.9 γίνεται ο έλεγχος παλινδρόμησης. Έτσι, οι τιμές των συντελεστών της εξίσωσης παλινδρόμησης είναι α = -0,3177, b = -1,461 και R = -0,99. Ο έλεγχος t-Student δείχνει ότι η χρήση του διαγράμματος ταξινόμησης δεν ενδείκνυται για όλο τον πληθυσμό.

6.19 Σενάριο 18°

Στο σενάριο αυτό, περιλαμβάνονται όλα τα ομαδοποιημένα χαρακτηριστικά, έχοντας προσθέσει την μεταβλητή Thickness:

Thickness
DIM1
DIM2
DIM3
DIM4
DIM5

- Στον πίνακα 18.1 παρατίθονται τα στατιστικά δεδομένα που προκύπτουν από τις δύο ομάδες, για κάθε χαρακτηριστικό χωριστά.
- Στην συνέχεια, η μαθηματική διαδικασία διακόπτεται γιατί δεν υπάρχει ο αντίστροφος πίνακας του W. Επομένως είναι αδύνατος ο προσδιορισμός της διαχωριστικής συνάρτησης και των άλλων παραμέτρων της ανάλυσης.

6.20 Σενάριο 19°

Στο σενάριο αυτό, περιλαμβάνονται όλα τα ομαδοποιημένα χαρακτηριστικά, έχοντας προσθέσει την μεταβλητή Pressure:

Pressure
DIM1
DIM2
DIM3
DIM4
DIM5

- Στον πίνακα 19.1 παρατίθονται τα στατιστικά δεδομένα που προκύπτουν από τις δύο ομάδες, για κάθε χαρακτηριστικό χωριστά.
- Στην συνέχεια, η μαθηματική διαδικασία διακόπτεται γιατί δεν υπάρχει ο αντίστροφος πίνακας του W. Επομένως είναι αδύνατος ο προσδιορισμός της διαχωριστικής συνάρτησης και των άλλων παραμέτρων της ανάλυσης.

6.21 Σενάριο 20°

Στο σενάριο αυτό, περιλαμβάνονται όλα τα ομαδοποιημένα χαρακτηριστικά, έχοντας προσθέσει την μεταβλητή Life-time:

Life-time
DIM1
DIM2
DIM3
DIM4
DIM5

- Στον πίνακα 20.1 παρατίθονται τα στατιστικά δεδομένα που προκύπτουν από τις δύο ομάδες, για κάθε χαρακτηριστικό χωριστά.

- Στον πίνακα 20.2 εμφανίζονται τα αποτελέσματα της διαχωριστικής ανάλυσης. Το χαρακτηριστικό με την μεγαλύτερη βαρύτητα και διαχωριστική ισχύ είναι το Life-time, ενώ αυτό με την μικρότερη είναι το DIM3. Επίσης, όλα τα χαρακτηριστικά είναι ευσταθή εκτός των DIM1 και DIM5 που παρουσίασαν αστάθεια.
- Στον πίνακα 20.3 παρουσιάζεται ο πίνακας ταξινόμησης, όπου φαίνονται τα ακόλουθα ποσοστά σωστής ταξινόμησης :

$$EF_1=91,7\%$$
$$EF_2=98,3\%$$
$$EF_t=94,0\%$$

- Το διάγραμμα ταξινόμησης φαίνεται στον πίνακα 20.5 και δείχνει τον καλό διαχωρισμό που επιτυγχάνεται μεταξύ των δύο ομάδων.
- Στον πίνακα 20.4 δίνεται η τιμή του Rao's F = 57,9943 για τον έλεγχο ισότητας των μέσων τιμών. Από τους πίνακες της κατανομής F προκύπτει ότι για όλα τα επίπεδα σημαντικότητας, οι μέσες τιμές της μεταβλητής Z των δύο ομάδων δεν είναι ίσες. Η τιμή του Wilk's Λ στην περίπτωση αυτή είναι 0,3163.
- Στον πίνακα 20.6 δίνεται η τιμή των Box's B = 362,2751 και Box's F = 16,4959 και γίνεται ο έλεγχος ισότητας των διασπορών.
- Από τον πίνακα 20.7 προκύπτει ότι ο Κανονικός Συντελεστής Συσχέτισης είναι C.C.C. = 0,8268. Η αντίστοιχη ιδιοτιμή ισούται με λ = 2,16128.
- Ο έλεγχος προσαρμογής κανονικής κατανομής της μεταβλητής Z, και για τις δύο ομάδες, φαίνεται στον πίνακα 20.8. Προκύπτει ότι σε όλα τα επίπεδα σημαντικότητας (α=1%, 5%, 10%) η μεταβλητή Z ακολουθεί την κανονική κατανομή.
- Τέλος, στον πίνακα 20.9 γίνεται ο έλεγχος παλινδρόμησης. Έτσι, οι τιμές των συντελεστών της εξίσωσης παλινδρόμησης είναι α = -1,3948, b = 3,0616 και R = 0,99. Ο έλεγχος t-Student δείχνει ότι η χρήση του διαγράμματος ταξινόμησης ενδείκνυται για όλο τον πληθυσμό.

6.22 Σενάριο 21°

Το σενάριο αυτό, περιλαμβάνει μεταβλητές, στις οποίες εμφανίζονται χαρακτηριστικά μόνο μια φορά, μεμονωμένα ή ομαδοποιημένα:

Thickness
Life-time
DIM1
DIM2

- Στον πίνακα 21.1 παρατίθονται τα στατιστικά δεδομένα που προκύπτουν από τις δύο ομάδες, για κάθε χαρακτηριστικό χωριστά.
- Στην συνέχεια, η μαθηματική διαδικασία διακόπτεται γιατί δεν υπάρχει ο αντίστροφος πίνακας του W - Within Groups SSCP. Επομένως είναι αδύνατος ο προσδιορισμός της διαχωριστικής συνάρτησης και των άλλων παραμέτρων της ανάλυσης.

6.23 Σενάριο 22°

Το σενάριο αυτό, περιλαμβάνει μεταβλητές, στις οποίες εμφανίζονται χαρακτηριστικά μόνο μια φορά, μεμονωμένα ή ομαδοποιημένα:

Length
Pressure
Life-time
DIM3

- Στον πίνακα 22.1 παρατίθονται τα στατιστικά δεδομένα που προκύπτουν από τις δύο ομάδες, για κάθε χαρακτηριστικό χωριστά.
- Στον πίνακα 22.2 εμφανίζονται τα αποτελέσματα της διαχωριστικής ανάλυσης. Το χαρακτηριστικό με την μεγαλύτερη βαρύτητα και διαχωριστική ισχύ

είναι το Life-time, ενώ αυτό με την μικρότερη είναι το Length. Επίσης, όλα τα χαρακτηριστικά είναι ευσταθή.

- Στον πίνακα 22.3 παρουσιάζεται ο πίνακας ταξινόμησης, όπου φαίνονται τα ακόλουθα ποσοστά σωστής ταξινόμησης :

$$EF_1=86,2\%$$
$$EF_2=94,9\%$$
$$EF_t=89,3\%$$

- Το διάγραμμα ταξινόμησης φαίνεται στον πίνακα 22.5 και δείχνει τον μέτριο διαχωρισμό που επιτυγχάνεται μεταξύ των δύο ομάδων.
- Στον πίνακα 22.4 δίνεται η τιμή του Rao's F = 52,3467 για τον έλεγχο ισότητας των μέσων τιμών. Από τους πίνακες της κατανομής F προκύπτει ότι για όλα τα επίπεδα σημαντικότητας, οι μέσες τιμές της μεταβλητής Z των δύο ομάδων δεν είναι ίσες. Επίσης η τιμή του Wilk's Λ = 0,4377 υποδεικνύει τον ίδιο μέτριο διαχωρισμό που επιτυγχάνεται μεταξύ των δύο ομάδων.
- Στον πίνακα 22.6 δίνεται η τιμή των Box's B = 115,389 και Box's F = 11,1985 και γίνεται ο έλεγχος ισότητας των διασπορών.
- Από τον πίνακα 22.7 προκύπτει ότι ο Κανονικός Συντελεστής Συσχέτισης είναι C.C.C. = 0,7499. Η αντίστοιχη ιδιοτιμή ισούται με λ = 1,28458.
- Ο έλεγχος προσαρμογής κανονικής κατανομής της μεταβλητής Z, και για τις δύο ομάδες, φαίνεται στον πίνακα 22.8. Προκύπτει ότι σε όλα τα επίπεδα σημαντικότητας (α=1%, 5%, 10%) η μεταβλητή Z ακολουθεί την κανονική κατανομή.
- Τέλος, στον πίνακα 22.9 γίνεται ο έλεγχος παλινδρόμησης. Έτσι, οι τιμές των συντελεστών της εξίσωσης παλινδρόμησης είναι α = -0,8290, b = -2,3602 και R = -0,99. Ο έλεγχος t-Student δείχνει ότι η χρήση του διαγράμματος ταξινόμησης δεν ενδείκνυται για όλο τον πληθυσμό.

Κεφάλαιο

7

ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΠΡΟΤΑΣΕΙΣ

Στα γενικά πλαίσια, ο σκοπός της εργασίας αυτής ήταν να μελετήσει την χρησιμότητα της εφαρμογής των πολυμεταβλητών στατιστικών τεχνικών της διαχωριστικής ανάλυσης και της ταξινόμησης, με την χρήση ομαδοποιημένων μεταβλητών, στην πρόβλεψη της αξιοπιστίας ενός συστήματος υποθαλάσσιων αγωγών μεταφοράς πετρελαίου στην Βόρεια Θάλασσα.

Ωστόσο, ο σκοπός αυτός αναδεικνύεται πολύπλευρος, καθώς δόθηκε ιδιαίτερη σημασία:

- στην χρησιμότητα των δύο στατιστικών τεχνικών, κυρίως με την συνδυαστική εφαρμογή τους,
- στην ανάλυση της συμπεριφοράς και επίδρασης του κάθε χαρακτηριστικού στην αξιοπιστία των αγωγών,
- στην λογική της δημιουργίας συνδυασμών χαρακτηριστικών με σκοπό την βελτίωση της στατιστικής ακρίβειας των αποτελεσμάτων,
- στην σημασία της εισαγωγής ομαδοποιημένων χαρακτηριστικών και στα αποτελέσματα που προέκυψαν από την ανάλυση αυτή,
- στην ιδιαιτερότητα της χρήσης **αδιάστατων** ομαδοποιημένων μεταβλητών,

- στην αλληλεξάρτηση και αλληλεπίδραση μεταξύ των απλών και των ομαδοποιημένων χαρακτηριστικών,
- ενδεχομένως, στον προσδιορισμό στοιχειωδών κανόνων ομαδοποίησης και επιλογής χαρακτηριστικών.

Από την μελέτη που προηγήθηκε, προκύπτουν μεταξύ άλλων και τα εξής συμπεράσματα:

- Οι πολυμεταβλητές στατιστικές τεχνικές της διαχωριστικής ανάλυσης και της ταξινόμησης, αποτελούν ένα πολύτιμο εργαλείο, χρήσιμο για την επίλυση προβλημάτων αξιοπιστίας και όχι μόνο.
- Το διάγραμμα ταξινόμησης, το οποίο αποτελεί το τελικό προϊόν του συνδυασμού των δύο τεχνικών, αποδεικνύεται απλό στην χρήση του ενώ παράλληλα αποτελεί κατάληξη θεμελιωμένων στατιστικών κανόνων και μαθηματικών σχέσεων.
- Η εισαγωγή **ομαδοποιημένων μεταβλητών** στην πρόβλεψη αξιοπιστίας συμβάλει, από κάθε άποψη, σε μεγάλο βαθμό στην δημιουργία μοντέλων πρόβλεψης τα οποία παρέχουν πολύ υψηλά ποσοστά ταξινόμησης. Το συμπέρασμα αυτό προκύπτει άλλωστε από το γεγονός ότι τα βέλτιστα μοντέλα, στο σύνολο αυτών που μελετήθηκαν, περιλαμβάνουν ομαδοποιημένες μεταβλητές.
- Αναμφισβήτητα, τα ομαδοποιημένα χαρακτηριστικά περιγράφουν με τον βέλτιστο τρόπο την αλληλεξάρτηση και αλληλεπίδραση μεταξύ των διαφόρων χαρακτηριστικών. Κατά την μελέτη των διαφόρων σεναρίων, παρατηρήθηκε το φαινόμενο της υποβάθμισης ή της αναβάθμισης της ισχύος ενός μοντέλου με την εισαγωγή ή απαλοιφή μιας ή περισσοτέρων ομαδοποιημένων μεταβλητών. Το γεγονός αυτό μπορεί να ερμηνευτεί ως εξής: Η εισαγωγή π.χ. μιας ομαδοποιημένης μεταβλητής μπορεί να έχει δύο αποτελέσματα:
  i. την εισαγωγή πληροφοριών που «φέρει» το ομαδοποιημένο χαρακτηριστικό στην διαχωριστική συνάρτηση, τα οποία αναδεικνύουν την διαχωριστική ισχύ ορισμένων άλλων χαρακτηριστικών και τονίζουν την βαρύ-

τητα της αλληλεπίδρασης μεταξύ διαφόρων μεταβλητών στην αξιοπιστία των αγωγών,

ii. την εισαγωγή πληροφοριών στην διαχωριστική συνάρτηση οι οποίες όμως ήδη προϋπάρχουν με αποτέλεσμα την υποβάθμιση της διαχωριστικής ισχύος του πλεονάζοντος χαρακτηριστικού.

Χαρακτηριστικό παράδειγμα αυτού, φαίνεται στο αποτέλεσμα της διαχωριστικής ανάλυσης που παρουσιάζεται στον πίνακα 20.2. Παρατηρείται δηλαδή το γεγονός, ότι ενώ η ομαδοποιημένη μεταβλητή DIM2 σε όλα τα μοντέλα που περιέχεται εμφανίζει πάντα την μεγαλύτερη διαχωριστική ισχύ, η εισαγωγή του χαρακτηριστικού Life-time (χρόνος ζωής) υποβαθμίζει την ισχύ της, θέτοντας την ίδια πρώτη στην κατάταξη, αυξάνοντας σημαντικά τα ποσοστά σωστής ταξινόμησης. Άλλωστε, μια πρόχειρη ματιά δείχνει τις σημαντικές εναλλαγές στην «ιεραρχία» των μεταβλητών από μοντέλο σε μοντέλο.

☞ Σαφώς δεν είναι τυχαία η τόσο υψηλή διαχωριστική ικανότητα που εμφανίζει η μεταβλητή DIM2. Φαίνεται να εκφράζει σε μεγάλο βαθμό την σχέση που υπάρχει ανάμεσα στο μήκος του αγωγού και στην διάμετρό του, και την επίδραση στην αξιοπιστία του συστήματος.

☞ Τα συνολικά αποτελέσματα φαίνεται να δείχνουν την σημαντικότητα της χρησιμοποίησης αδιάστατων ομαδοποιημένων χαρακτηριστικών. Ιδιαίτερο άλλωστε ενδιαφέρον έχει το γεγονός ότι μέσω αυτών εκφράζονται οι επιδράσεις τρίτων παραγόντων και φαινομένων που μπορούν να επηρεάζουν την αξιοπιστία του συστήματος. Ωστόσο, τα αποτελέσματα δείχνουν να επηρεάζουν κατά περίπτωση την ισχύ ενός μοντέλου. Ως παράδειγμα τούτου, αναφέρεται η χαρακτηριστική σύγκριση[1] μεταξύ των μεταβλητών:

$$A = \frac{25{,}4 \cdot L \cdot D}{(0{,}1 \cdot T)^2} \qquad B = \frac{25{,}4 \cdot L \cdot D}{(0{,}1 \cdot T)^{-2}}$$

Έτσι, ενώ η αδιάστατη μεταβλητή A παρουσιάζει πολύ ασθενή διαχωριστική ισχύ, η μεταβλητή B [m^4] εμφανίζεται, απεναντίας, αρκετά «ισχυ-

ρή». Το γεγονός αυτό δείχνει ότι είναι πολύ δύσκολο με εμπειρικούς τρόπους να προσδιορισθούν εκείνοι οι συνδυασμοί που θα βελτιστοποιούν την δημιουργία αδιάστατων, και γενικότερα ομαδοποιημένων, μεταβλητών.

- Η αύξηση του αριθμού των χαρακτηριστικών, οι τιμές των οποίων λαμβάνονται υπόψη, έχει σαν αποτέλεσμα τον καλύτερο διαχωρισμό και την αύξηση των ποσοστών σωστής ταξινόμησης. Οδηγεί, δηλαδή, σχεδόν πάντα στην δημιουργία ενός βελτιωμένου μοντέλου πρόβλεψης της συμπεριφοράς των αγωγών.
- Η σημασία του προσήμου στους συντελεστές της διαχωριστικής συνάρτησης (πίνακας 1.2, 2.2, κ.ο.κ.) σχετίζεται κάθε φορά με τον τρόπο που επηρεάζει το συγκεκριμένο χαρακτηριστικό την τιμή της διαχωριστικής μεταβλητής. Έτσι, π.χ. αρνητικός συντελεστής ενός χαρακτηριστικού Χ δηλώνει ότι αύξηση της τιμής του χαρακτηριστικού αυτού έχει ως συνέπεια την μείωση της τιμής της μεταβλητής Z και επομένως αυξάνεται η πιθανότητα ταξινόμησης του αγωγού στην ομάδα των επιτυχιών, και αντίστροφα. Εντούτοις, η ερμηνεία του προσήμου πρέπει να γίνεται κατά περίπτωση σε συνάρτηση με τα άλλα χαρακτηριστικά που περιλαμβάνονται στην ανάλυση.
- Λαμβάνοντας υπόψη το σύνολο των στατιστικών παραμέτρων, το καλύτερο σενάριο πρόβλεψης προκύπτει το 10°, με πολύ μικρή διαφορά από το 1° και το 8°. Στο σενάριο αυτό περιλαμβάνονται όλα τα χαρακτηριστικά, εκτός από το DIM3, το οποίο στις περισσότερες περιπτώσεις εμφανίζει πολύ μικρή διαχωριστική ικανότητα. Τα ποσοστά σωστής ταξινόμησης έχουν ως εξής: Για τις επιτυχίες EF_1 = 95,4%, για τις αποτυχίες EF_2 = 98,3% και συνολικά EF_t = 96,4%. Η τιμή του Wilk's Λ είναι ίση με 0,2796 και ο κανονικός συντελεστής συσχέτισης ισούται με C.C.C. = 0,8487.
- Συνολικά, τα μοντέλα που παρουσίασαν τα καλύτερα αποτελέσματα ήταν τα: 1°, 4°, 8°, 10°, 11°, 12°, 14° και 20°, ενώ ως το χειρότερο χαρακτηρίζεται το 15°. Σε όλες αυτές τις περιπτώσεις τα ποσοστά σωστής ταξινόμησης κυμαίνονται σε επίπεδα άνω του 90%. Οι τιμές του Wilk's Λ κυμαίνονται από 0,27 ÷ 0,30, ενώ οι τιμές του C.C.C. γύρω στο 0,84.

[1] Βλέπε Y. L. Bakouros, «Offshore Pipeline Reliability Prediction», University of Bradford, 1988, σελίδα 141-142.

Όλα τα παραπάνω μας οδηγούν στο τελικό, πλέον, συμπέρασμα της μεγάλης χρησιμότητας και σημασίας της εφαρμογής των πολυμεταβλητών στατιστικών τεχνικών στην πρόβλεψη της αξιοπιστίας. Ιδιαίτερη, επίσης, έμφαση δίνεται στην χρήση ομαδοποιημένων μεταβλητών ως μέσο βελτιστοποίησης των δύο τεχνικών, της διαχωριστικής ανάλυσης και της ταξινόμησης. Σε ένα μελλοντικό στάδιο διεύρυνσης της παρούσας μελέτης, θα ήταν δυνατή η εφαρμογή των τεχνικών αυτών για την πρόβλεψη της αξιοπιστίας και άλλων μηχανικών συστημάτων ή εξαρτημάτων. Ενδιαφέροντα συμπεράσματα αναμένεται να προκύψουν από την σύγκριση των αποτελεσμάτων της πρακτικής εφαρμογής των διαδικασιών αυτών.

Τέλος, ο συγγραφέας ελπίζει η μελέτη που περιλαμβάνεται στο παρόν σύγγραμμα να αποτελέσει αφορμή οποιασδήποτε περαιτέρω έρευνας, με απώτερο στόχο την εμβάθυνση στην γνώση και την κατανόηση της περιπλοκότητας των αλληλεπιδράσεων μεταξύ των μεταβλητών, οι οποίες επηρεάζουν τον τρόπο με τον οποίο επέρχεται η αστοχία των αγωγών.

ΒΙΒΛΙΟΓΡΑΦΙΑ

- Andrews, G. 1984. Offshore Pipeline availability. Offshore Oil & Gas Pipeline Technology. 1984 European Seminar N.E.C.
- Armitage, P., McPherson, C.K. and Copas, J.C. 1969. Statistical studies of prognosis in advanced breast cancer. Journal of Chronic Diseases. Vol.22. pp. 343-360.
- Bakouros, Y. L., Offshore Pipeline Reliability Prediction, διδακτορική διατριβή, University of Bradford, 1988.
- Betts, J. 1984. The identification of companies at risk of financial failure. University of Bradford, Postgraduate school of studies in Industrial Technology. PhD thesis. Unpublished.
- Box, G.E.P. 1949. A general distribution theory for a class of likehood criteria. Biometrica. Vol. 36. pp. 317-346.
- Bulbrook, R.D., Hayward, J.L. and Thomas, B.S. 1964. The relation between the urinary 17-hydroxycorticosteroids and U-deoxy-17oxosteroids and the fate of patients after mestectomy. Lancet. i. pp. 945-947.
- Cannon, A.J. and Lewis, R.C. 1985. The reliability of pipe systems operating in the British sector of the North Sea. 1985. Paper 4A/R. Reliability Conference Proccedings N.E.C.
- Cox, D.R. and Brandwood, L. 1959. On a discriminatory problem connected with the works of Plato. Journal of the Royal Statistical Society. Series B. 21. pp. 195-200.
- de la Mare, R.F. and Anderson, O. 1980. Pipeline reliability. Det Norske Veritas. Report No. 80-0572.
- de la Mare, R.F. and Bakouros, Y.L. 1987. Pipeline Reliability Prediction using Discriminant Analysis. 1987 PPP Conference Proccedings. Utrecht.

- Det Norske Veritas. 1980. Report No. 80-0082. Failures and Failure Rates in Various Pipeline «Networks». T. Andersen.
- Fisher, R.A. 1936. The use of multiple measurements in taxonomic problems. Anuals of Eugenies 7 : 179-198.
- Johnson R. And Wichern D., «Applied multivariate statistical analysis», second edition, Prentice-Hall International Editions, London, 1988.
- Klecka, W.R. 1973. The clientele of Australian Parties. New perspectives through discriminant analysis. Politics 7. pp. 301-308.
- Rao, C.R. 1965. Linear statistical inference and its applications. New York. John Wiley.
- Sayles, R.S. 1980. The use of discriminant functions and factors analysis techniques in reliability assesment and data lassification. 6th Advances in Reliability Technology Symposium. University of Bradford.
- Veritec 1984. Paper No. 1984-002. Reliability and economics of pipeline operations. O. Furnes, T. Andersen.
- Tatsuoka, M.M. and Tieteman, D.V. 1954. Discriminant Analysis. Review of Educational Research. Vol. 24. pp. 402-420.
- Tatsuoka, M.M. 1971. Multivariate analysis. New York. John Wiley.
- Βατσάλη, Α. και Κωνσταντούλη, Κ. 1992. Μοντέλο διαχωριστικής ανάλυσης και ταξινόμησης με χρήση Η/Υ. Διπλωματική εργασία. Θεσσαλονίκη.
- Καπετανάκης, Γ.Ι. και Τσομίδου Α.Δ. 1995. Βελτιστοποίηση πρόβλεψης αξιοπιστίας με την χρήση ομαδοποιημένων μεταβλητών. Διπλωματική εργασία. Θεσσαλονίκη.
- Μπακούρος, Ι.Λ. 1992. Θεωρία Αξιοπιστίας. Θεσσαλονίκη.
- Ψωϊνός, Δ. Εφαρμοσμένη Στατιστική. Τρίτη έκδοση. Εκδόσεις Ζήτη. Θεσσαλονίκη.

ΠΑΡΑΡΤΗΜΑ

ΠΙΝΑΚΕΣ ΚΑΙ ΔΙΑΓΡΑΜΜΑΤΑ ΔΙΑΧΩΡΙΣΤΙΚΗΣ ΑΝΑΛΥΣΗΣ & ΤΑΞΙΝΟΜΗΣΗΣ

ΜΕΤΑΒΛΗΤΕΣ	ΜΕΣΗ ΤΙΜΗ	ΤΥΠ.ΑΠΟΚ/ΣΗ	ΜΕΓΙΣΤΟ	ΕΛΑΧΙΣΤΟ
LENGTH....	50.358	63.897	452.000	1.000
DIAMETER..	23.734	20.108	255.800	3.000
THICKNESS.	0.611	0.166	1.000	0.079
PRESSURE..	112.633	31.498	255.000	18.000
LIFETIME..	64.720	37.543	221.000	0.000
DIM1.....	41.773	24.566	230.000	14.353
DIM2.....	48.416	57.169	427.160	0.328
DIM3.....	30.589	12.921	83.333	16.778
DIM4.....	13.574	7.812	34.668	0.484
DIM5.....	6.281	11.657	56.121	0.000

Πίνακας 1.1

ΜΕΤΑΒΛΗΤΕΣ	Ui	Si	ΚΑΤΑΤΑΞΗ	ΣΤΑΘΕΡΟΤΗΤΑ	Uo
LENGTH....	0.009384	0.599630	3	ΕΥΣΤΑΘΗΣ	1.771887
DIAMETER..	0.000203	0.004079	10	ΕΥΣΤΑΘΗΣ	
THICKNESS.	-2.281368	-0.378179	5	ΑΣΤΑΘΗΣ	
PRESSURE..	0.008501	0.267760	7	ΕΥΣΤΑΘΗΣ	
LIFETIME..	-0.026448	-0.992927	2	ΕΥΣΤΑΘΗΣ	
DIM1.....	0.011562	0.284019	6	ΑΣΤΑΘΗΣ	
DIM2.....	-0.018529	-1.059289	1	ΕΥΣΤΑΘΗΣ	
DIM3.....	-0.008218	-0.106182	9	ΕΥΣΤΑΘΗΣ	
DIM4.....	0.019779	0.154508	8	ΕΥΣΤΑΘΗΣ	
DIM5.....	0.047045	0.548389	4	ΑΣΤΑΘΗΣ	

Πίνακας 1.2

Π Ι Ν Α Κ Α Σ Κ Α Ι Π Ο Σ Ο Σ Τ Α Τ Α Ξ Ι Ν Ο Μ Η Σ Η Σ

Ποσοστά Σωστής Ταξινόμησης

ΕΠΙΤΥΧΙΕΣ .... EF1 = 95.4%
ΑΠΟΤΥΧΙΕΣ .... EF2 = 96.6%
ΣΥΝΟΛΙΚΑ ..... EFt = 95.8%

		ΠΡΟΒΛΕΠΟΜΕΝΑ: ΕΠΙΤΥΧΙΕΣ	ΠΡΟΒΛΕΠΟΜΕΝΑ: ΑΠΟΤΥΧΙΕΣ	
ΠΡΑΓΜΑΤΙΚΑ	ΕΠΙΤΥΧΙΕΣ	104	5	109
	ΑΠΟΤΥΧΙΕΣ	2	57	59
		106	62	

ESCAPE : Επιστροφή

Πίνακας 1.3

≡≡≡ Επιτυχίες Αποτυχίες Στατιστική Ανάλυση Εξοδος

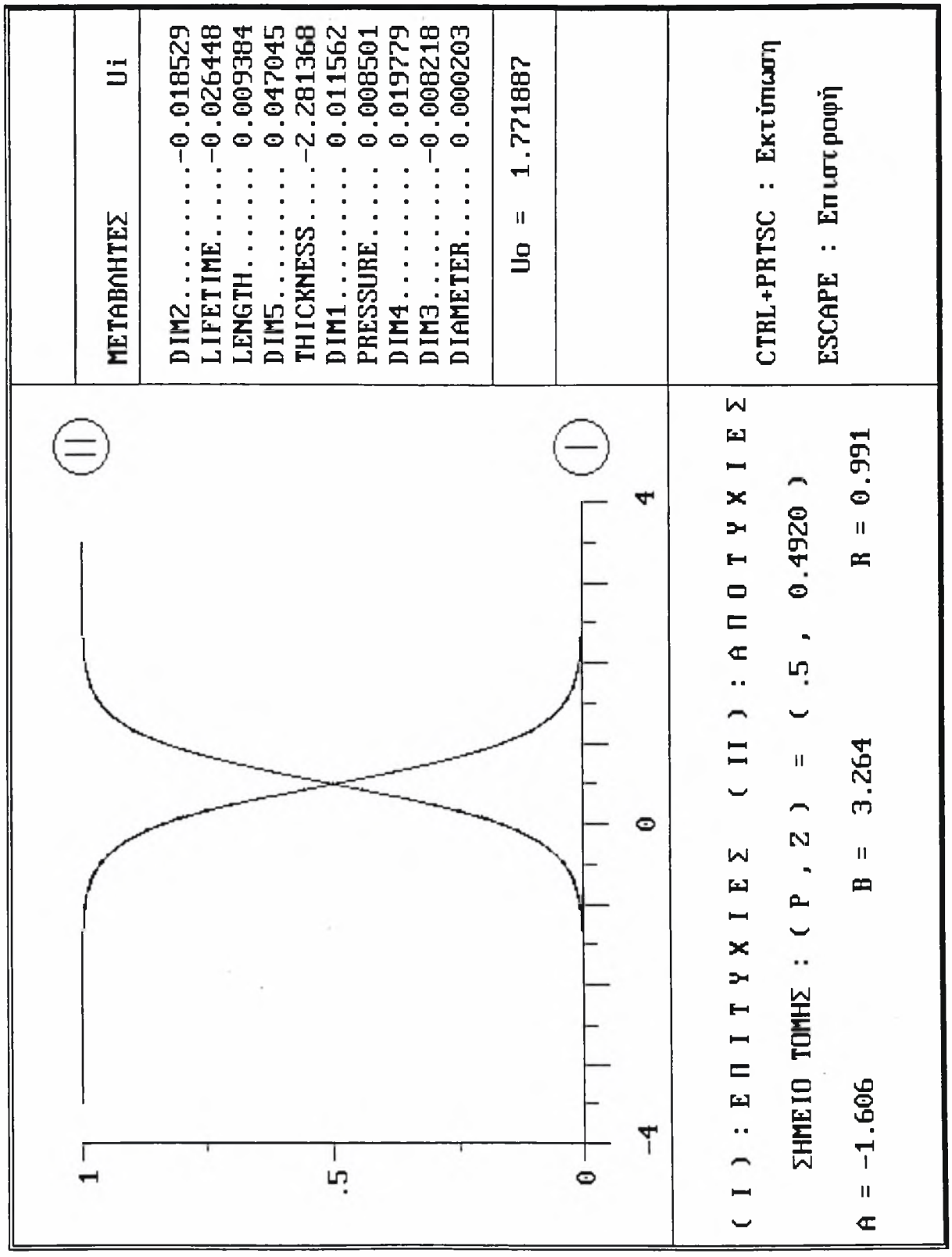
Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
ης Z

R A O ' S F			
Wilk's Λ	F	v1	v2
0.2788	40.6177	10	157

Πάτησε οποιοδήποτε πλήκτρο για επιστροφή

Πίνακας 1.4



Πίνακας 1.5

≡≡≡	Επιτυχίες	Αποτυχίες	Στατιστική Ανάλυση	Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
ης Z

B O X ' S F			
Box's B	F	v1	v2
1149.3682	19.3888	55	47777
Πάτησε οποιοδήποτε πλήκτρο για επιστροφή			

Πίνακας 1.6

≡≡≡	Επιτυχίες	Αποτυχίες	Στατιστική Ανάλυση	Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

σων τιμών
ασπορών
ς συσχέτισης
τα της Z
ση

Ιδιοτιμές	C.C.C
0.00000	0.0000
2.58712	0.8492
-0.00000	0.0000
-0.00000	0.0000
0.00000	0.0004
0.00000	0.0003
-0.00000	0.0000
0.00000	0.0002
0.00000	0.0002
0.00000	0.0000
Πάτησε οποιοδήποτε πλήκτρο	

Πίνακας 1.7

≡≡≡ Επιτυχίες Αποτυχίες Στατιστική Ανάλυση Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών

Ε λ ε γ χ ο ς K o l m o g o r o v - S m i r n o v		
	Ε π ι τ υ χ ί ε ς	Α π ο τ υ χ ί ε ς
Μέση τιμή	-1.176305	2.173175
Μεταβλητότητα	1.119385	.7455422
Επ. σημαντικότητας	Ολα τα επίπεδα	Ολα τα επίπεδα

Πίνακας 1.8

≡≡≡ Επιτυχίες Αποτυχίες Στατιστική Ανάλυση Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
Κανονικότητα της Z
Παλινδρόμηση

Η εξίσωση παλινδρόμησης στις επιτυχίες έχει :
a = -1.6057 b = 3.2637 R = 0.9915
Αφού έγινε ο έλεγχος κατά Student προέκυψε
για τον πληθυσμό R > 0.

Πάτησε οποιοδήποτε πλήκτρο για συνέχεια.

Πίνακας 1.9

ΜΕΤΑΒΛΗΤΕΣ	ΜΕΣΗ ΤΙΜΗ	ΤΥΠ.ΑΠΟΚ/ΣΗ	ΜΕΓΙΣΤΟ	ΕΛΑΧΙΣΤΟ
LENGTH....	50.358	63.897	452.000	1.000
DIAMETER..	23.734	20.108	255.800	3.000
THICKNESS.	0.611	0.166	1.000	0.079
PRESSURE..	112.633	31.498	255.000	18.000
LIFETIME..	64.720	37.543	221.000	0.000

Πίνακας 2.1

ΜΕΤΑΒΛΗΤΕΣ	Ui	Si	ΚΑΤΑΤΑΞΗ	ΣΤΑΘΕΡΟΤΗΤΑ	Uo
LENGTH....	-0.004584	-0.292928	2	ΕΥΣΤΑΘΗΣ	-1.317716
DIAMETER..	-0.006815	-0.137033	3	ΕΥΣΤΑΘΗΣ	
THICKNESS.	0.649512	0.107669	4	ΑΣΤΑΘΗΣ	
PRESSURE..	-0.003093	-0.097419	5	ΕΥΣΤΑΘΗΣ	
LIFETIME..	0.025679	0.964066	1	ΕΥΣΤΑΘΗΣ	

Πίνακας 2.2

Π Ι Ν Α Κ Α Σ Κ Α Ι Π Ο Σ Ο Σ Τ Α Τ Α Ξ Ι Ν Ο Μ Η Σ Η Σ

Ποσοστά Σωστής Ταξινόμησης

ΕΠΙΤΥΧΙΕΣ .... EF1 = 82.6%
ΑΠΟΤΥΧΙΕΣ .... EF2 = 96.6%
ΣΥΝΟΛΙΚΑ ..... EFt = 87.5%

		ΠΡΟΒΛΕΠΟΜΕΝΑ: ΕΠΙΤΥΧΙΕΣ	ΠΡΟΒΛΕΠΟΜΕΝΑ: ΑΠΟΤΥΧΙΕΣ	
ΠΡΑΓΜΑΤΙΚΑ	ΕΠΙΤΥΧΙΕΣ	90	19	109
	ΑΠΟΤΥΧΙΕΣ	2	57	59
		92	76	

ESCAPE : Επιστροφή

Πίνακας 2.3

≡≡≡ Επιτυχίες Αποτυχίες Στατιστική Ανάλυση Εξοδος

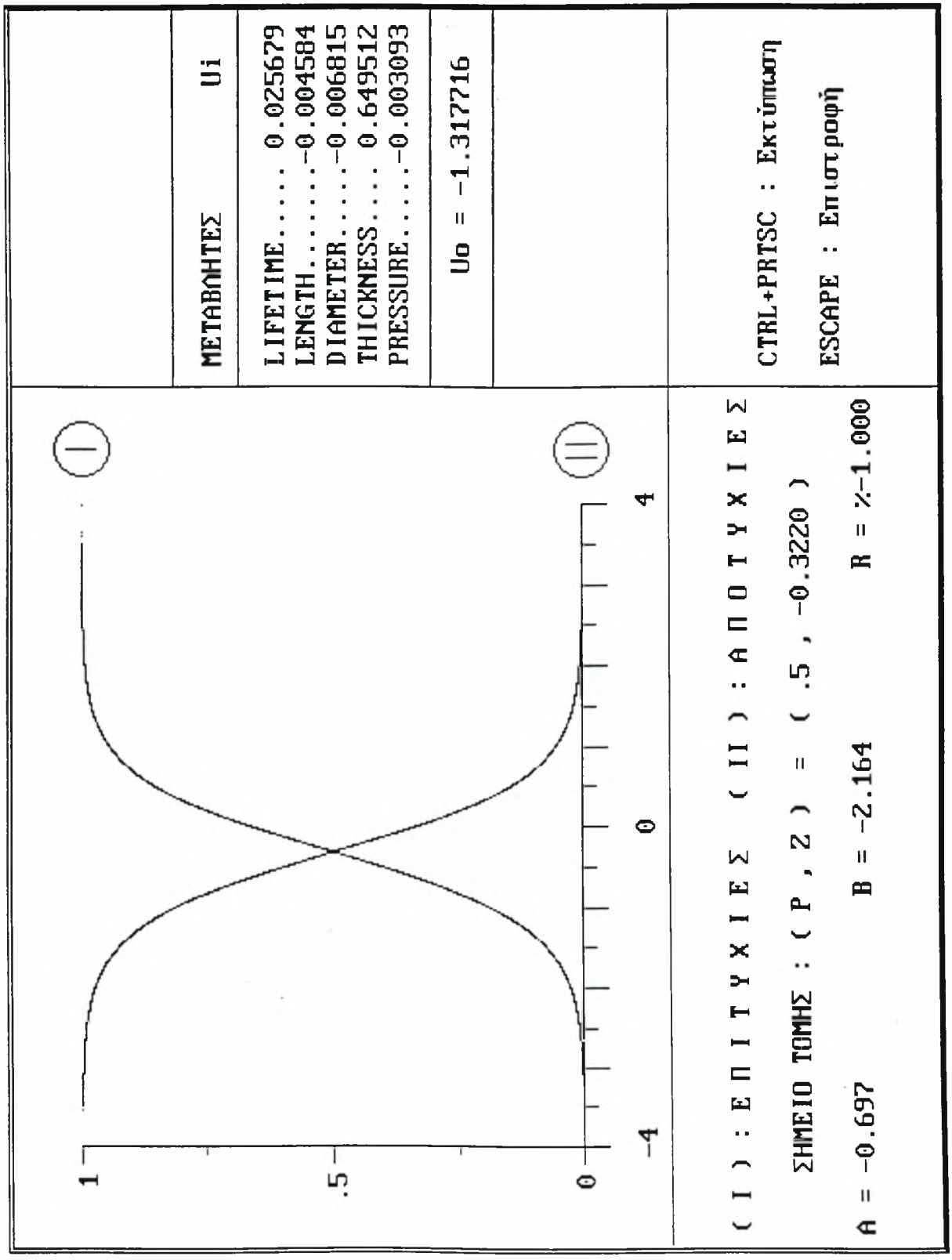
Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
ης Z

R A O ' S F			
Wilk's Λ	F	v1	v2
0.4808	34.9822	5	162

Πάτησε οποιοδήποτε πλήκτρο για επιστροφή

Πίνακας 2.4



Πίνακας 2.5

≡≡≡	Επιτυχίες	Αποτυχίες	Στατιστική Ανάλυση	Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
ης Z

B O X ' S F			
Box's B	F	v1	v2
209.6393	13.4636	15	58057
Πάτησε οποιοδήποτε πλήκτρο για επιστροφή			

Πίνακας 2.6

≡≡≡	Επιτυχίες	Αποτυχίες	Στατιστική Ανάλυση	Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

σων τιμών
ασπορών
ς συσχέτισης
τα της Z
ση

Ιδιοτιμές	C.C.C
0.00000	0.0000
1.07970	0.7205
0.00000	0.0002
-0.00000	0.0000
-0.00000	0.0000
Πάτησε οποιοδήποτε πλήκτρο	

Πίνακας 2.7

```
===    Επιτυχίες    Αποτυχίες    Στατιστική Ανάλυση    Εξοδος
                                  Συνολικά στατιστικά στοιχεία
                                  Διαχωριστική ανάλυση
                                  Πίνακας ταξινόμησης
                                  ξινόμησης
                                  κασίας
                Ισότητα μέσων τιμών
                Ισότητα διασπορών
```

Ε λ ε γ χ ο ς	K o l m o g o r o v - S m i r n o v	
	Ε π ι τ υ χ ί ε ς	Α π ο τ υ χ ί ε ς
Μέση τιμή	.7599122	-1.403905
Μεταβλητότητα	1.273528	.4607705
Επ. σημαντικότητας	Ολα τα επίπεδα	Ολα τα επίπεδα

Πίνακας 2.8

```
===    Επιτυχίες    Αποτυχίες    Στατιστική Ανάλυση    Εξοδος
                                  Συνολικά στατιστικά στοιχεία
                                  Διαχωριστική ανάλυση
                                  Πίνακας ταξινόμησης
                                  ξινόμησης
                                  κασίας
                Ισότητα μέσων τιμών
                Ισότητα διασπορών
                Συντελεστης συσχέτισης
                Κανονικότητα της Z
                Παλινδρόμηση

Η εξίσωση παλινδρόμησης στις επιτυχίες έχει :
 a = -0.6967     b = -2.1638     R = -0.9900
 Αφού έγινε ο έλεγχος κατά Student προέκυψε
 για τον πληθυσμό R = 0.

   Πάτησε οποιοδήποτε πλήκτρο για συνέχεια.
```

Πίνακας 2.9

ΜΕΤΑΒΛΗΤΕΣ	ΜΕΣΗ ΤΙΜΗ	ΤΥΠ.ΑΠΟΚ/ΣΗ	ΜΕΓΙΣΤΟ	ΕΛΑΧΙΣΤΟ
LENGTH....	50.358	63.897	452.000	1.000
DIAMETER..	23.734	20.108	255.800	3.000
THICKNESS.	0.611	0.166	1.000	0.079
PRESSURE..	112.633	31.498	255.000	18.000
LIFETIME..	64.720	37.543	221.000	0.000
DIM1.....	41.773	24.566	230.000	14.353

Πίνακας 3.1

ΜΕΤΑΒΛΗΤΕΣ	Ui	Si	ΚΑΤΑΤΑΞΗ	ΣΤΑΘΕΡΟΤΗΤΑ	Uo
LENGTH....	0.004399	0.281054	4	ΕΥΣΤΑΘΗΣ	0.063392
DIAMETER..	0.007264	0.146069	5	ΕΥΣΤΑΘΗΣ	
THICKNESS.	-0.721700	-0.119635	6	ΑΣΤΑΘΗΣ	
PRESSURE..	0.010205	0.321449	2	ΕΥΣΤΑΘΗΣ	
LIFETIME..	-0.026280	-0.986630	1	ΕΥΣΤΑΘΗΣ	
DIM1.....	0.012805	0.314569	3	ΑΣΤΑΘΗΣ	

Πίνακας 3.2

Π Ι Ν Α Κ Α Σ Κ Α Ι Π Ο Σ Ο Σ Τ Α Τ Α Ξ Ι Ν Ο Μ Η Σ Η Σ

Ποσοστά Σωστής Ταξινόμησης

ΕΠΙΤΥΧΙΕΣ .... EF1 = 82.6%
ΑΠΟΤΥΧΙΕΣ .... EF2 = 94.9%
ΣΥΝΟΛΙΚΑ ..... EFt = 86.9%

		ΠΡΟΒΛΕΠΟΜΕΝΑ: ΕΠΙΤΥΧΙΕΣ	ΠΡΟΒΛΕΠΟΜΕΝΑ: ΑΠΟΤΥΧΙΕΣ	
ΠΡΑΓΜΑΤΙΚΑ	ΕΠΙΤΥΧΙΕΣ	90	19	109
	ΑΠΟΤΥΧΙΕΣ	3	56	59
		93	75	

ESCAPE : Επιστροφή

Πίνακας 3.3

≡≡≡ Επιτυχίες Αποτυχίες Στατιστική Ανάλυση Εξοδος

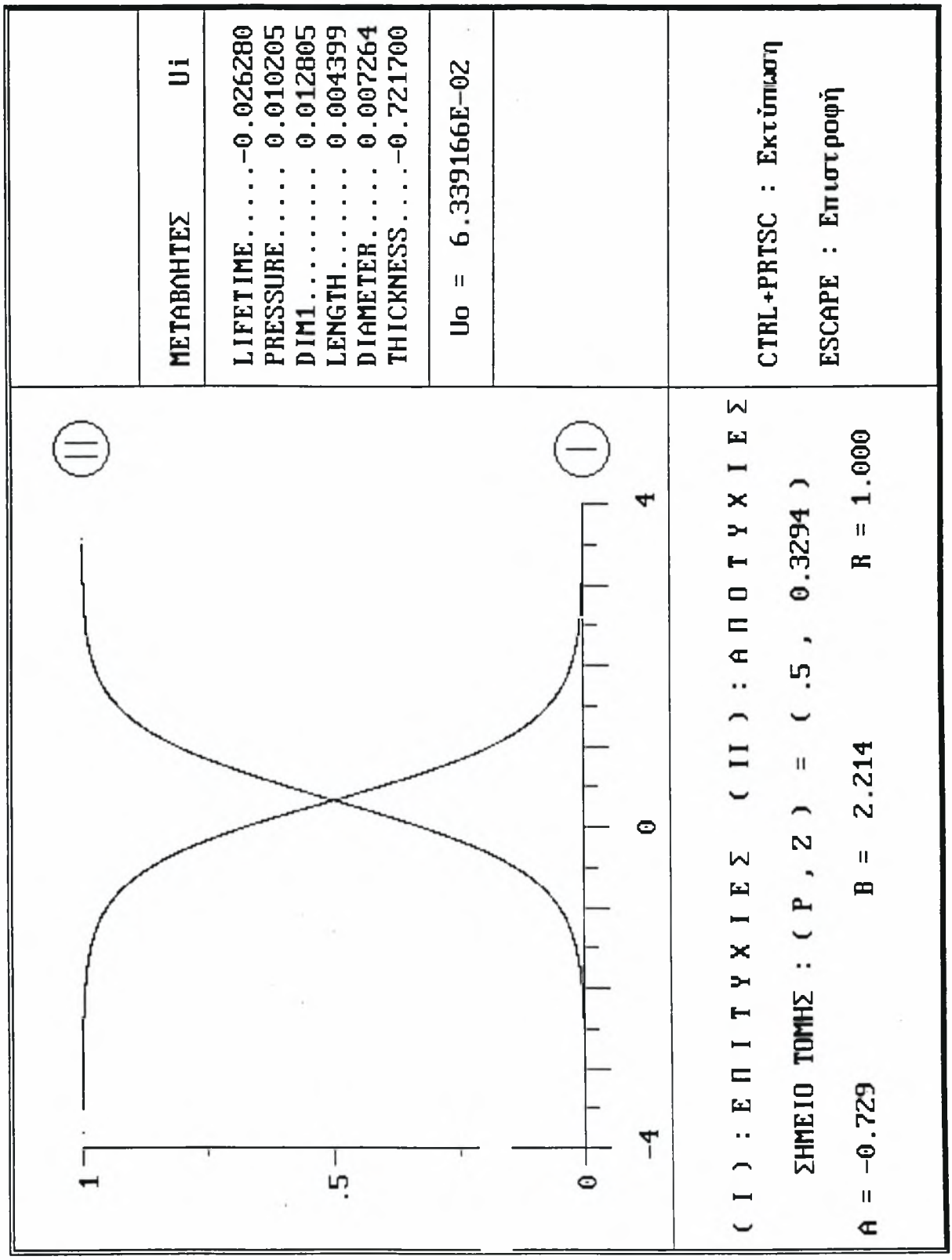
Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
ης Z

R A O ' S F			
Wilk's Λ	F	v1	v2
0.4695	30.3209	6	161

Πάτησε οποιοδήποτε πλήκτρο για επιστροφή

Πίνακας 3.4



Πίνακας 3.5

≡≡≡	Επιτυχίες	Αποτυχίες	Στατιστική Ανάλυση	Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
ης Z

B O X ' S F			
Box's B	F	v1	v2
315.3169	14.3577	21	53659
Πάτησε οποιοδήποτε πλήκτρο για επιστροφή			

Πίνακας 3.6

≡≡≡	Επιτυχίες	Αποτυχίες	Στατιστική Ανάλυση	Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

σων τιμών
ασπορών
ς συσχέτισης
τα της Z
ση

Ιδιοτιμές	C.C.C
0.00000	0.0000
1.12997	0.7284
0.00000	0.0004
0.00000	0.0002
-0.00000	0.0000
-0.00000	0.0000
Πάτησε οποιοδήποτε πλήκτρο	

Πίνακας 3.7

```
===   Επιτυχίες   Αποτυχίες   Στατιστική Ανάλυση   Εξοδος

                              Συνολικά στατιστικά στοιχεία
                              Διαχωριστική ανάλυση
                              Πίνακας ταξινόμησης
                              ξινόμησης
                              κασίας
                 Ισότητα μέσων τιμών
                 Ισότητα διασπορών
```

Ε λ ε γ χ ο ς	K o l m o g o r o v - S m i r n o v	
	Ε π ι τ υ χ ί ε ς	Α π ο τ υ χ ί ε ς
Μέση τιμή	-.7774024	1.436218
Μεταβλητότητα	1.252873	.4989296
Επ. σημαντικότητας	Ολα τα επίπεδα	Ολα τα επίπεδα

Πίνακας 3.8

```
===   Επιτυχίες   Αποτυχίες   Στατιστική Ανάλυση   Εξοδος

                              Συνολικά στατιστικά στοιχεία
                              Διαχωριστική ανάλυση
                              Πίνακας ταξινόμησης
                              ξινόμησης
                              κασίας
                 Ισότητα μέσων τιμών
                 Ισότητα διασπορών
                 Συντελεστης συσχέτισης
                 Κανονικότητα της Z
                 Παλινδρόμηση

Η εξίσωση παλινδρόμησης στις επιτυχίες έχει :
 a = -0.7292     b =  2.2136     R =  1.0000
 Αφού έγινε ο έλεγχος κατά Student προέκυψε
 για τον πληθυσμό R > 0.

   Πάτησε οποιοδήποτε πλήκτρο για συνέχεια.
```

Πίνακας 3.9

ΜΕΤΑΒΛΗΤΕΣ	ΜΕΣΗ ΤΙΜΗ	ΤΥΠ.ΑΠΟΚ/ΣΗ	ΜΕΓΙΣΤΟ	ΕΛΑΧΙΣΤΟ
LENGTH....	50.358	63.897	452.000	1.000
DIAMETER..	23.734	20.108	255.800	3.000
THICKNESS.	0.611	0.166	1.000	0.079
PRESSURE..	112.633	31.498	255.000	18.000
LIFETIME..	64.720	37.543	221.000	0.000
DIM2.....	48.416	57.169	427.160	0.328

Πίνακας 4.1

ΜΕΤΑΒΛΗΤΕΣ	Ui	Si	ΚΑΤΑΤΑΞΗ	ΣΤΑΘΕΡΟΤΗΤΑ	Uo
LENGTH....	-0.013999	-0.894497	2	ΕΥΣΤΑΘΗΣ	-1.830509
DIAMETER..	-0.003577	-0.071916	6	ΕΥΣΤΑΘΗΣ	
THICKNESS.	2.077566	0.344395	4	ΑΣΤΑΘΗΣ	
PRESSURE..	-0.006148	-0.193655	5	ΕΥΣΤΑΘΗΣ	
LIFETIME..	0.019212	0.721290	3	ΕΥΣΤΑΘΗΣ	
DIM2.....	0.016532	0.945096	1	ΕΥΣΤΑΘΗΣ	

Πίνακας 4.2

Π Ι Ν Α Κ Α Σ Κ Α Ι Π Ο Σ Ο Σ Τ Α Τ Α Ξ Ι Ν Ο Μ Η Σ Η Σ

Ποσοστά Σωστής Ταξινόμησης

ΕΠΙΤΥΧΙΕΣ .... EF1 = 95.4%
ΑΠΟΤΥΧΙΕΣ .... EF2 = 94.9%
ΣΥΝΟΛΙΚΑ ..... EFt = 95.2%

		ΠΡΟΒΛΕΠΟΜΕΝΑ: ΕΠΙΤΥΧΙΕΣ	ΠΡΟΒΛΕΠΟΜΕΝΑ: ΑΠΟΤΥΧΙΕΣ	
ΠΡΑΓΜΑΤΙΚΑ	ΕΠΙΤΥΧΙΕΣ	104	5	109
	ΑΠΟΤΥΧΙΕΣ	3	56	59
		107	61	

ESCAPE : Επιστροφή

Πίνακας 4.3

=== Επιτυχίες Αποτυχίες Στατιστική Ανάλυση Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

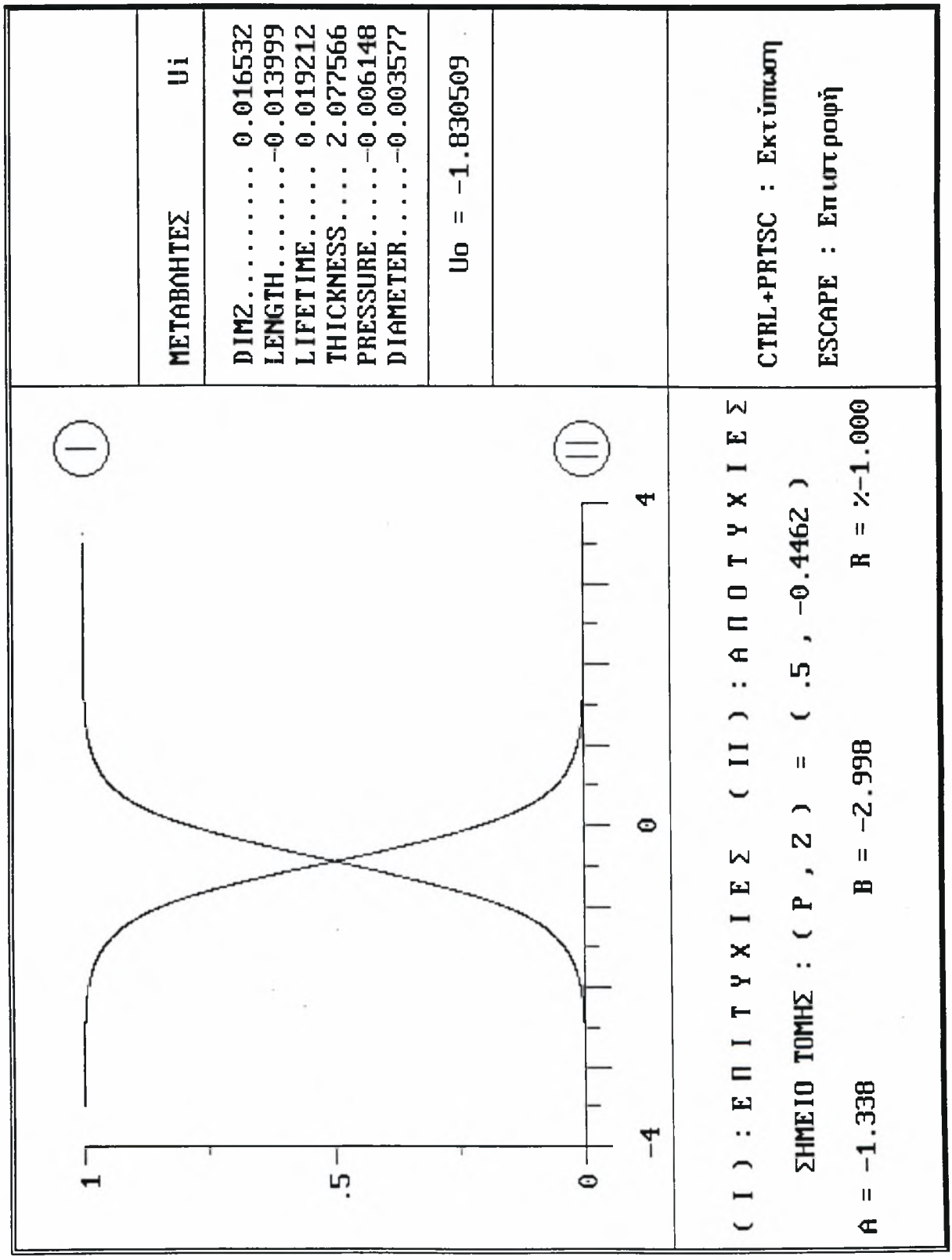
Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
ης Z

R A O ' S F

Wilk's Λ	F	v1	v2
0.3254	55.6321	6	161

Πάτησε οποιοδήποτε πλήκτρο για επιστροφή

Πίνακας 4.4



Πίνακας 4.5

≡≡≡	Επιτυχίες	Αποτυχίες	Στατιστική Ανάλυση	Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
ης Z

B O X ' S F			
Box's B	F	v1	v2
668.0142	30.4175	21	53659
Πάτησε οποιοδήποτε πλήκτρο για επιστροφή			

Πίνακας 4.6

≡≡≡	Επιτυχίες	Αποτυχίες	Στατιστική Ανάλυση	Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

σων τιμών
ασπορών
ς συσχέτισης
τα της Z
ση

Ιδιοτιμές	C.C.C
0.00000	0.0000
2.07325	0.8213
-0.00000	0.0000
0.00000	0.0001
-0.00000	0.0000
0.00000	0.0002
Πάτησε οποιοδήποτε πλήκτρο	

Πίνακας 4.7

```
===   Επιτυχίες        Αποτυχίες        Στατιστική Ανάλυση        Εξοδος
                                          | Συνολικά στατιστικά στοιχεία
                                          | Διαχωριστική ανάλυση
                                          | Πίνακας ταξινόμησης
                                          | ξινόμησης
                                          | κασίας
                  Ισότητα μέσων τιμών
                  Ισότητα διασπορών
```

Ε λ ε γ χ ο ς	K o l m o g o r o v - S m i r n o v	
	Ε π ι τ υ χ ί ε ς	Α π ο τ υ χ ί ε ς
Μέση τιμή	1.053022	-1.945413
Μεταβλητότητα	.9877309	.9887673
Επ. σημαντικότητας	Ολα τα επίπεδα	Ολα τα επίπεδα

Πίνακας 4.8

```
===   Επιτυχίες        Αποτυχίες        Στατιστική Ανάλυση        Εξοδος
                                          | Συνολικά στατιστικά στοιχεία
                                          | Διαχωριστική ανάλυση
                                          | Πίνακας ταξινόμησης
                                          | ξινόμησης
                                          | κασίας
                  Ισότητα μέσων τιμών
                  Ισότητα διασπορών
                  Συντελεστης συσχέτισης
                  Κανονικότητα της Z
                  Παλινδρόμηση

Η εξίσωση παλινδρόμησης στις επιτυχίες έχει :
 a = -1.3379      b = -2.9984      R = -0.9900
 Αφού έγινε ο έλεγχος κατά Student προέκυψε
 για τον πληθυσμό R = 0.

   Πάτησε οποιοδήποτε πλήκτρο για συνέχεια.
```

Πίνακας 4.9

ΜΕΤΑΒΛΗΤΕΣ	ΜΕΣΗ ΤΙΜΗ	ΤΥΠ.ΑΠΟΚ/ΣΗ	ΜΕΓΙΣΤΟ	ΕΛΑΧΙΣΤΟ
LENGTH....	50.358	63.897	452.000	1.000
DIAMETER..	23.734	20.108	255.800	3.000
THICKNESS.	0.611	0.166	1.000	0.079
PRESSURE..	112.633	31.498	255.000	18.000
LIFETIME..	64.720	37.543	221.000	0.000
DIM3.....	30.589	12.921	83.333	16.778

Πίνακας 5.1

ΜΕΤΑΒΛΗΤΕΣ	Ui	Si	ΚΑΤΑΤΑΞΗ	ΣΤΑΘΕΡΟΤΗΤΑ	Uo
LENGTH....	-0.002306	-0.147346	4	ΕΥΣΤΑΘΗΣ	-2.129401
DIAMETER..	-0.000620	-0.012465	6	ΕΥΣΤΑΘΗΣ	
THICKNESS.	0.282764	0.046873	5	ΑΣΤΑΘΗΣ	
PRESSURE..	-0.006115	-0.192617	3	ΕΥΣΤΑΘΗΣ	
LIFETIME..	0.025997	0.976013	1	ΕΥΣΤΑΘΗΣ	
DIM3.....	0.035756	0.462001	2	ΕΥΣΤΑΘΗΣ	

Πίνακας 5.2

Π Ι Ν Α Κ Α Σ Κ Α Ι Π Ο Σ Ο Σ Τ Α Τ Α Ξ Ι Ν Ο Μ Η Σ Η Σ

Ποσοστά Σωστής Ταξινόμησης

ΕΠΙΤΥΧΙΕΣ .... EF1 = 86.2%
ΑΠΟΤΥΧΙΕΣ .... EF2 = 94.9%
ΣΥΝΟΛΙΚΑ ..... EFt = 89.3%

		ΠΡΟΒΛΕΠΟΜΕΝΑ: ΕΠΙΤΥΧΙΕΣ	ΠΡΟΒΛΕΠΟΜΕΝΑ: ΑΠΟΤΥΧΙΕΣ	
ΠΡΑΓΜΑΤΙΚΑ	ΕΠΙΤΥΧΙΕΣ	94	15	109
	ΑΠΟΤΥΧΙΕΣ	3	56	59
		97	71	

ESCAPE : Επιστροφή

Πίνακας 5.3

=== Επιτυχίες Αποτυχίες Στατιστική Ανάλυση Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

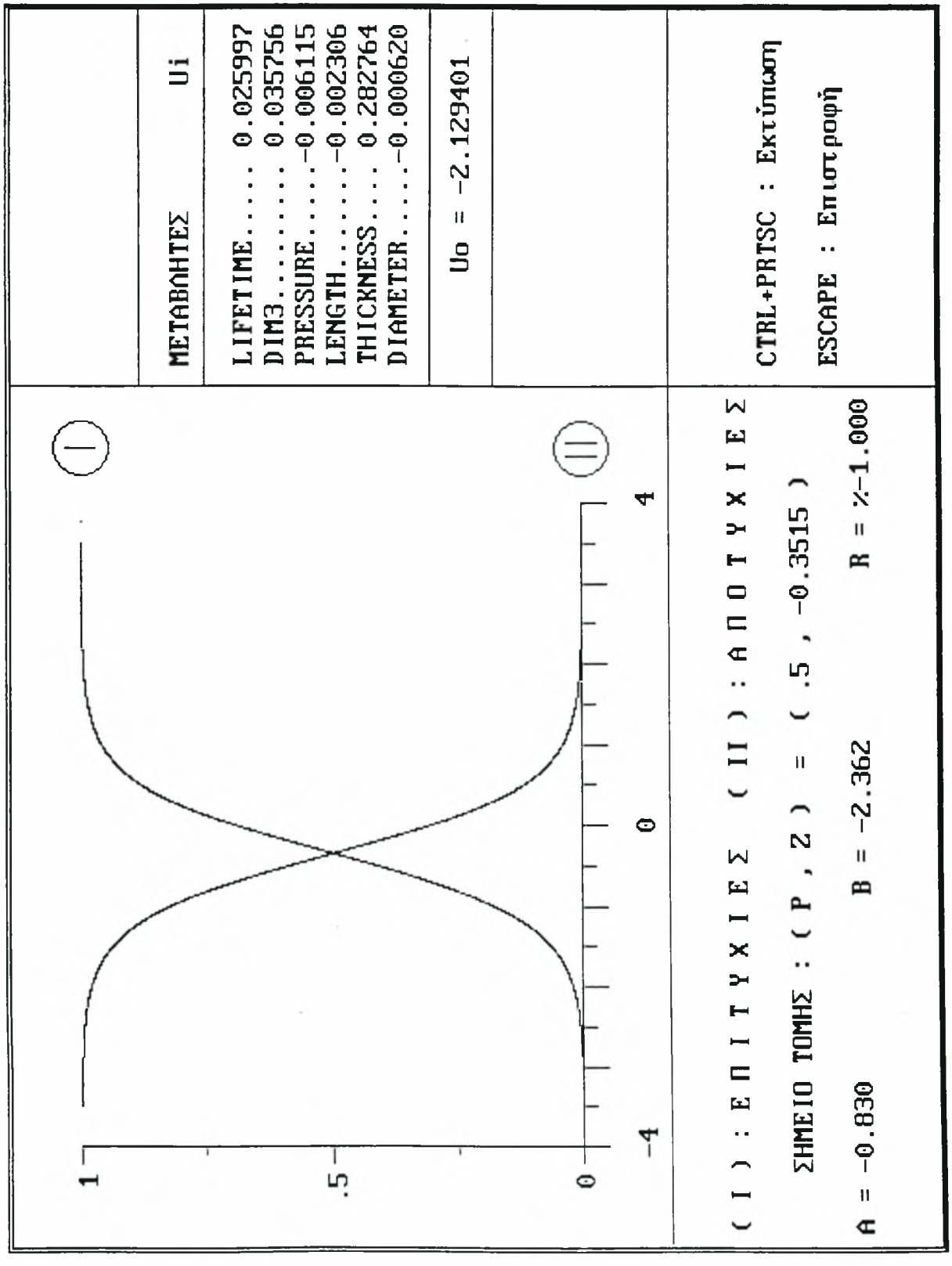
Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
ης Z

R A O ' S F

Wilk's Λ	F	v1	v2
0.4373	34.5247	6	161

Πάτησε οποιοδήποτε πλήκτρο για επιστροφή

Πίνακας 5.4



Πίνακας 5.5

≡≡≡ Επιτυχίες Αποτυχίες Στατιστική Ανάλυση Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
ης Z

B O X ' S F			
Box's B	F	v1	v2
375.2672	17.0875	21	53659

Πάτησε οποιοδήποτε πλήκτρο για επιστροφή

Πίνακας 5.6

≡≡≡ Επιτυχίες Αποτυχίες Στατιστική Ανάλυση Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

σων τιμών
ασπορών
ς συσχέτισης
τα της Z
ση

Ιδιοτιμές	C.C.C
-4.46438	0.0000
1.28663	0.7501
0.00000	0.0005
0.00000	0.0000
0.00000	0.0000
-0.00000	0.0000

Πάτησε οποιοδήποτε πλήκτρο

Πίνακας 5.7

≡≡≡	Επιτυχίες	Αποτυχίες	Στατιστική Ανάλυση	Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών

Ε λ ε γ χ ο ς	K o l m o g o r o v - S m i r n o v	
	Ε π ι τ υ χ ί ε ς	Α π ο τ υ χ ί ε ς
Μέση τιμή	.8295446	-1.532548
Μεταβλητότητα	1.23399	.5338141
Επ. σημαντικότητας	Ολα τα επίπεδα	Ολα τα επίπεδα

Πίνακας 5.8

≡≡≡	Επιτυχίες	Αποτυχίες	Στατιστική Ανάλυση	Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
Κανονικότητα της Z
Παλινδρόμηση

Η εξίσωση παλινδρόμησης στις επιτυχίες έχει :
a = -0.8303 b = -2.3621 R = -0.9890
Αφού έγινε ο έλεγχος κατά Student προέκυψε
για τον πληθυσμό R = 0.

Πάτησε οποιοδήποτε πλήκτρο για συνέχεια.

Πίνακας 5.9

ΜΕΤΑΒΛΗΤΕΣ	ΜΕΣΗ ΤΙΜΗ	ΤΥΠ.ΑΠΟΚ/ΣΗ	ΜΕΓΙΣΤΟ	ΕΛΑΧΙΣΤΟ
LENGTH....	50.358	63.897	452.000	1.000
DIAMETER..	23.734	20.108	255.800	3.000
THICKNESS.	0.611	0.166	1.000	0.079
PRESSURE..	112.633	31.498	255.000	18.000
LIFETIME..	64.720	37.543	221.000	0.000
DIM4.....	13.574	7.812	34.668	0.484

Πίνακας 6.1

ΜΕΤΑΒΛΗΤΕΣ	Ui	Si	ΚΑΤΑΤΑΞΗ	ΣΤΑΘΕΡΟΤΗΤΑ	Uo
LENGTH....	0.001156	0.073845	4	ΕΥΣΤΑΘΗΣ	1.506369
DIAMETER..	0.001703	0.034248	5	ΕΥΣΤΑΘΗΣ	
THICKNESS.	-0.169217	-0.028051	6	ΑΣΤΑΘΗΣ	
PRESSURE..	-0.007091	-0.223370	3	ΑΣΤΑΘΗΣ	
LIFETIME..	-0.027114	-1.017932	1	ΕΥΣΤΑΘΗΣ	
DIM4.....	0.077494	0.605371	2	ΕΥΣΤΑΘΗΣ	

Πίνακας 6.2

Π Ι Ν Α Κ Α Σ Κ Α Ι Π Ο Σ Ο Σ Τ Α Τ Α Ξ Ι Ν Ο Μ Η Σ Η Σ

Ποσοστά Σωστής Ταξινόμησης

ΕΠΙΤΥΧΙΕΣ .... EF1 = 89.0%
ΑΠΟΤΥΧΙΕΣ .... EF2 = 98.3%
ΣΥΝΟΛΙΚΑ ..... EFt = 92.3%

		ΠΡΟΒΛΕΠΟΜΕΝΑ: ΕΠΙΤΥΧΙΕΣ	ΠΡΟΒΛΕΠΟΜΕΝΑ: ΑΠΟΤΥΧΙΕΣ	
ΠΡΑΓΜΑΤΙΚΑ	ΕΠΙΤΥΧΙΕΣ	97	12	109
	ΑΠΟΤΥΧΙΕΣ	1	58	59
		98	70	

ESCAPE : Επιστροφή

Πίνακας 6.3

=== Επιτυχίες Αποτυχίες Στατιστική Ανάλυση Εξοδος

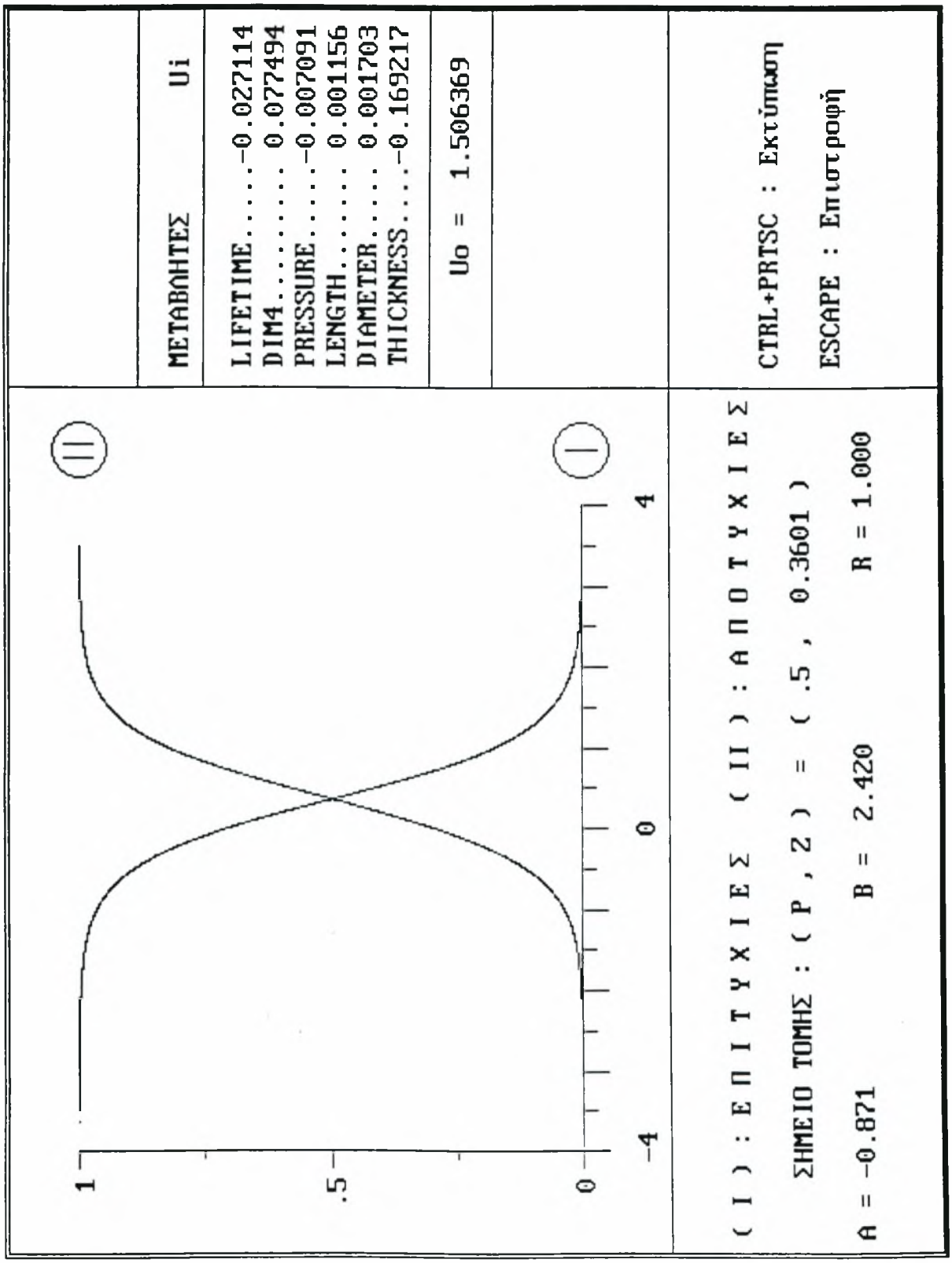
Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστής συσχέτισης
ης Z

R A O ' S F			
Wilk's Λ	F	v1	v2
0.4255	36.2315	6	161

Πάτησε οποιοδήποτε πλήκτρο για επιστροφή

Πίνακας 6.4



Πίνακας 6.5

≡≡≡	Επιτυχίες	Αποτυχίες	Στατιστική Ανάλυση	Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
ης Z

B O X ' S F			
Box's B	F	v1	v2
261.6633	11.9146	21	53659
Πάτησε οποιοδήποτε πλήκτρο για επιστροφή			

Πίνακας 6.6

≡≡≡	Επιτυχίες	Αποτυχίες	Στατιστική Ανάλυση	Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

σων τιμών
ασπορών
ς συσχέτισης
τα της Z
ση

Ιδιοτιμές	C.C.C
0.00000	0.0000
0.00000	0.0000
0.00000	0.0002
-0.00000	0.0000
1.35024	0.7580
0.00000	0.0001
Πάτησε οποιοδήποτε πλήκτρο	

Πίνακας 6.7

≡≡≡ Επιτυχίες Αποτυχίες Στατιστική Ανάλυση Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών

Ε λ ε γ χ ο ς	K o l m o g o r o v - S m i r n o v	
	Ε π ι τ υ χ ί ε ς	Α π ο τ υ χ ί ε ς
Μέση τιμή	-.849803	1.569974
Μεταβλητότητα	1.195233	.6054164
Επ. σημαντικότητας	Ολα τα επίπεδα	Ολα τα επίπεδα

Πίνακας 6.8

≡≡≡ Επιτυχίες Αποτυχίες Στατιστική Ανάλυση Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
Κανονικότητα της Z
Παλινδρόμηση

Η εξίσωση παλινδρόμησης στις επιτυχίες έχει :
a = -0.8713 b = 2.4198 R = 1.0000
Αφού έγινε ο έλεγχος κατά Student προέκυψε
για τον πληθυσμό R > 0.

Πάτησε οποιοδήποτε πλήκτρο για συνέχεια.

Πίνακας 6.9

ΜΕΤΑΒΛΗΤΕΣ	ΜΕΣΗ ΤΙΜΗ	ΤΥΠ.ΑΠΟΚ/ΣΗ	ΜΕΓΙΣΤΟ	ΕΛΑΧΙΣΤΟ
LENGTH....	50.358	63.897	452.000	1.000
DIAMETER..	23.734	20.108	255.800	3.000
THICKNESS.	0.611	0.166	1.000	0.079
PRESSURE..	112.633	31.498	255.000	18.000
LIFETIME..	64.720	37.543	221.000	0.000
DIM5.....	6.281	11.657	56.121	0.000

Πίνακας 7.1

ΜΕΤΑΒΛΗΤΕΣ	Ui	Si	ΚΑΤΑΤΑΞΗ	ΣΤΑΘΕΡΟΤΗΤΑ	Uo
LENGTH....	0.003985	0.254636	2	ΕΥΣΤΑΘΗΣ	1.418786
DIAMETER..	0.006692	0.134557	3	ΕΥΣΤΑΘΗΣ	
THICKNESS.	-0.670326	-0.111119	4	ΑΣΤΑΘΗΣ	
PRESSURE..	0.002729	0.085947	5	ΕΥΣΤΑΘΗΣ	
LIFETIME..	-0.026529	-0.995980	1	ΕΥΣΤΑΘΗΣ	
DIM5.....	0.006493	0.075687	6	ΑΣΤΑΘΗΣ	

Πίνακας 7.2

Π Ι Ν Α Κ Α Σ Κ Α Ι Π Ο Σ Ο Σ Τ Α Τ Α Ξ Ι Ν Ο Μ Η Σ Η Σ

Ποσοστά Σωστής Ταξινόμησης

ΕΠΙΤΥΧΙΕΣ .... EF1 = 82.6%
ΑΠΟΤΥΧΙΕΣ .... EF2 = 96.6%
ΣΥΝΟΛΙΚΑ ..... EFt = 87.5%

		ΠΡΟΒΛΕΠΟΜΕΝΑ: ΕΠΙΤΥΧΙΕΣ	ΠΡΟΒΛΕΠΟΜΕΝΑ: ΑΠΟΤΥΧΙΕΣ	
ΠΡΑΓΜΑΤΙΚΑ	ΕΠΙΤΥΧΙΕΣ	90	19	109
	ΑΠΟΤΥΧΙΕΣ	2	57	59
		92	76	

ESCAPE : Επιστροφή

Πίνακας 7.3

≡≡≡ Επιτυχίες Αποτυχίες Στατιστική Ανάλυση Εξοδο

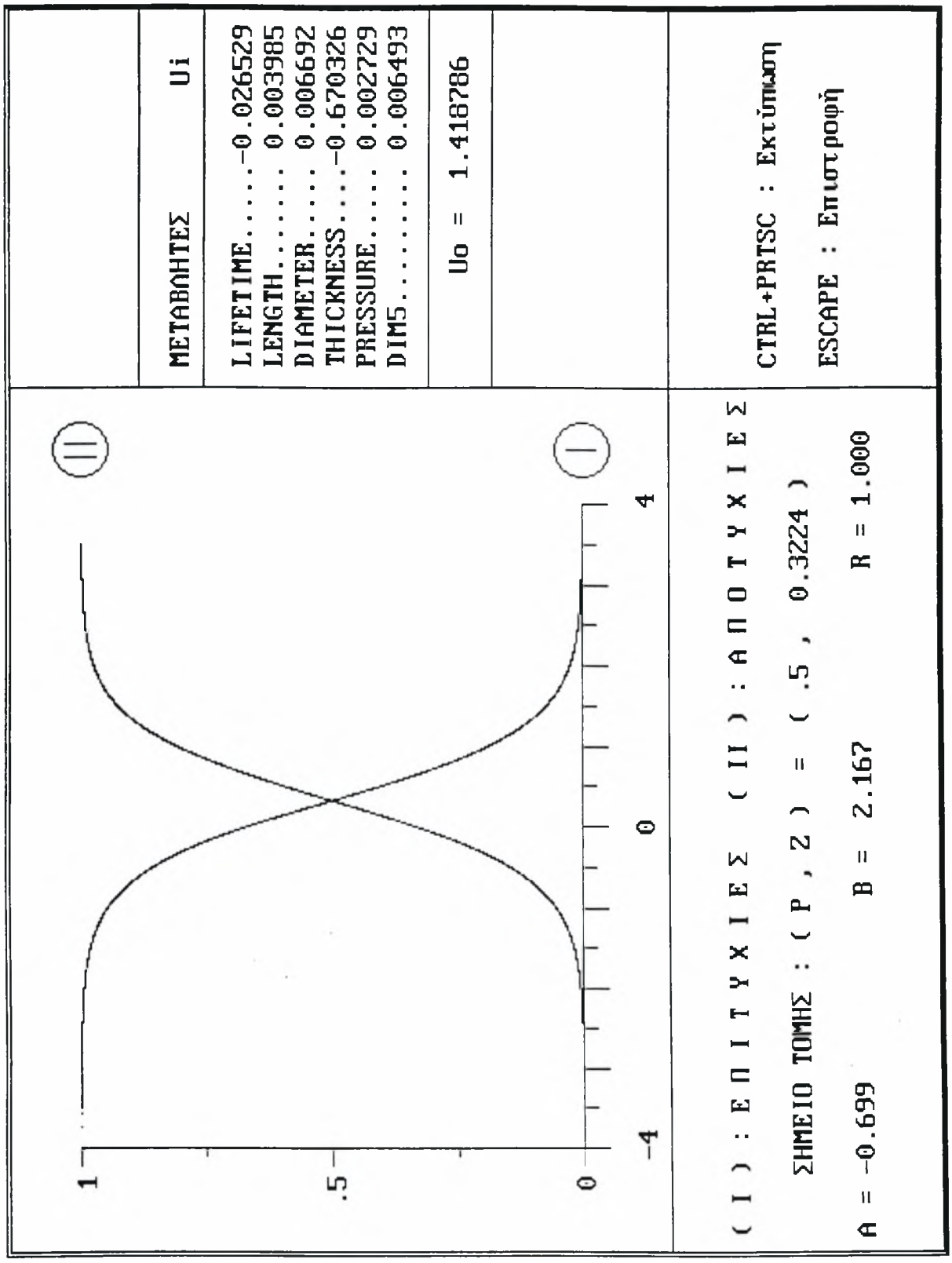
Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
ης Z

R A O ' S F			
Wilk's Λ	F	v1	v2
0.4802	29.0479	6	161

Πάτησε οποιοδήποτε πλήκτρο για επιστροφή

Πίνακας 7.4



Πίνακας 7.5

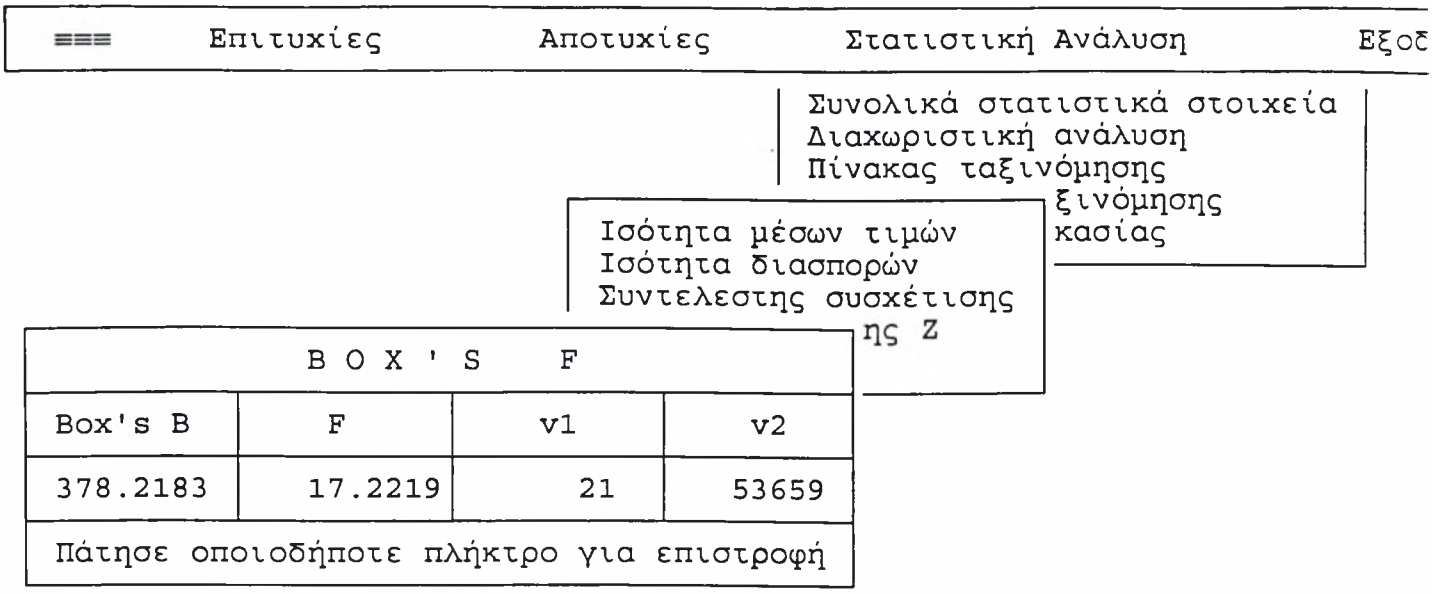
```
===    Επιτυχίες    Αποτυχίες    Στατιστική Ανάλυση    Εξοδος

                                 Συνολικά στατιστικά στοιχεία
                                 Διαχωριστική ανάλυση
                                 Πίνακας ταξινόμησης
                                 ξινόμησης
                                 κασίας
                Ισότητα μέσων τιμών
                Ισότητα διασπορών
```

Ε λ ε γ χ ο ς K o l m o g o r o v - S m i r n o v		
	Ε π ι τ υ χ ί ε ς	Α π ο τ υ χ ί ε ς
Μέση τιμή	-.760909	1.405747
Μεταβλητότητα	1.288663	.4328091
Επ. σημαντικότητας	Ολα τα επίπεδα	Ολα τα επίπεδα

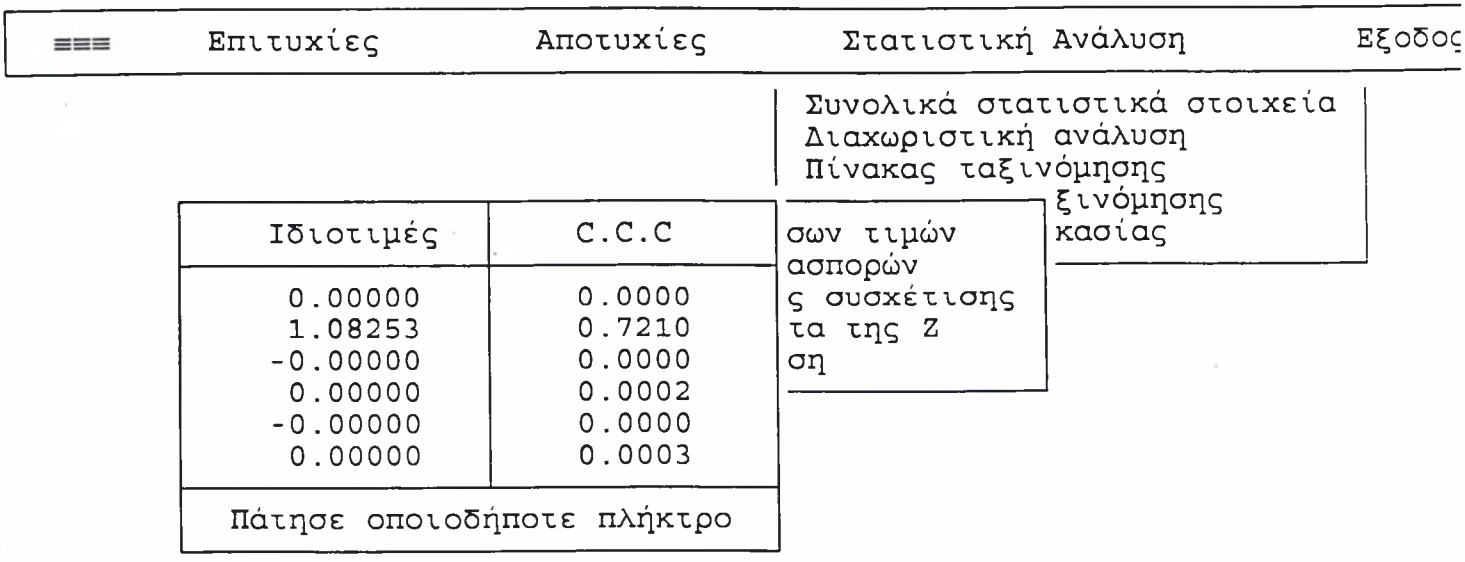
Πίνακας 7.6

```
===    Επιτυχίες    Αποτυχίες    Στατιστική Ανάλυση    Εξοδος

                                 Συνολικά στατιστικά στοιχεία
                                 Διαχωριστική ανάλυση
                                 Πίνακας ταξινόμησης
                                 ξινόμησης
                                 κασίας
                Ισότητα μέσων τιμών
                Ισότητα διασπορών
                Συντελεστης συσχέτισης
                Κανονικότητα της Z
                Παλινδρόμηση

Η εξίσωση παλινδρόμησης στις επιτυχίες έχει :
 a = -0.6986     b =  2.1667     R =  1.0000
 Αφού έγινε ο έλεγχος κατά Student προέκυψε
 για τον πληθυσμό R > 0.

   Πάτησε οποιοδήποτε πλήκτρο για συνέχεια.
```

Πίνακας 7.7



Πίνακας 7.8



Πίνακας 7.9

ΜΕΤΑΒΛΗΤΕΣ	ΜΕΣΗ ΤΙΜΗ	ΤΥΠ.ΑΠΟΚ/ΣΗ	ΜΕΓΙΣΤΟ	ΕΛΑΧΙΣΤΟ
LENGTH....	50.358	63.897	452.000	1.000
DIAMETER..	23.734	20.108	255.800	3.000
THICKNESS.	0.611	0.166	1.000	0.079
PRESSURE..	112.633	31.498	255.000	18.000
LIFETIME..	64.720	37.543	221.000	0.000
DIM2.....	48.416	57.169	427.160	0.328
DIM3.....	30.589	12.921	83.333	16.778
DIM4.....	13.574	7.812	34.668	0.484
DIM5.....	6.281	11.657	56.121	0.000

Πίνακας 8.1

ΜΕΤΑΒΛΗΤΕΣ	Ui	Si	ΚΑΤΑΤΑΞΗ	ΣΤΑΘΕΡΟΤΗΤΑ	Uo
LENGTH....	0.009871	0.630712	3	ΕΥΣΤΑΘΗΣ	
DIAMETER..	-0.000214	-0.004293	9	ΑΣΤΑΘΗΣ	
THICKNESS.	-2.300471	-0.381346	5	ΑΣΤΑΘΗΣ	
PRESSURE..	0.003843	0.121046	7	ΕΥΣΤΑΘΗΣ	
LIFETIME..	-0.025638	-0.962526	2	ΕΥΣΤΑΘΗΣ	2.996474
DIM2.....	-0.019028	-1.087818	1	ΕΥΣΤΑΘΗΣ	
DIM3.....	-0.011607	-0.149978	6	ΕΥΣΤΑΘΗΣ	
DIM4.....	0.007959	0.062177	8	ΕΥΣΤΑΘΗΣ	
DIM5.....	0.049581	0.577951	4	ΑΣΤΑΘΗΣ	

Πίνακας 8.2

Π Ι Ν Α Κ Α Σ Κ Α Ι Π Ο Σ Ο Σ Τ Α Τ Α Ξ Ι Ν Ο Μ Η Σ Η Σ

Ποσοστά Σωστής Ταξινόμησης

ΕΠΙΤΥΧΙΕΣ .... EF1 = 96.3%
ΑΠΟΤΥΧΙΕΣ .... EF2 = 96.6%
ΣΥΝΟΛΙΚΑ ..... EFt = 96.4%

		ΠΡΟΒΛΕΠΟΜΕΝΑ: ΕΠΙΤΥΧΙΕΣ	ΠΡΟΒΛΕΠΟΜΕΝΑ: ΑΠΟΤΥΧΙΕΣ	
ΠΡΑΓΜΑΤΙΚΑ	ΕΠΙΤΥΧΙΕΣ	105	4	109
	ΑΠΟΤΥΧΙΕΣ	2	57	59
		107	61	

ESCAPE : Επιστροφή

Πίνακας 8.3

=== Επιτυχίες Αποτυχίες Στατιστική Ανάλυση Εξοδος

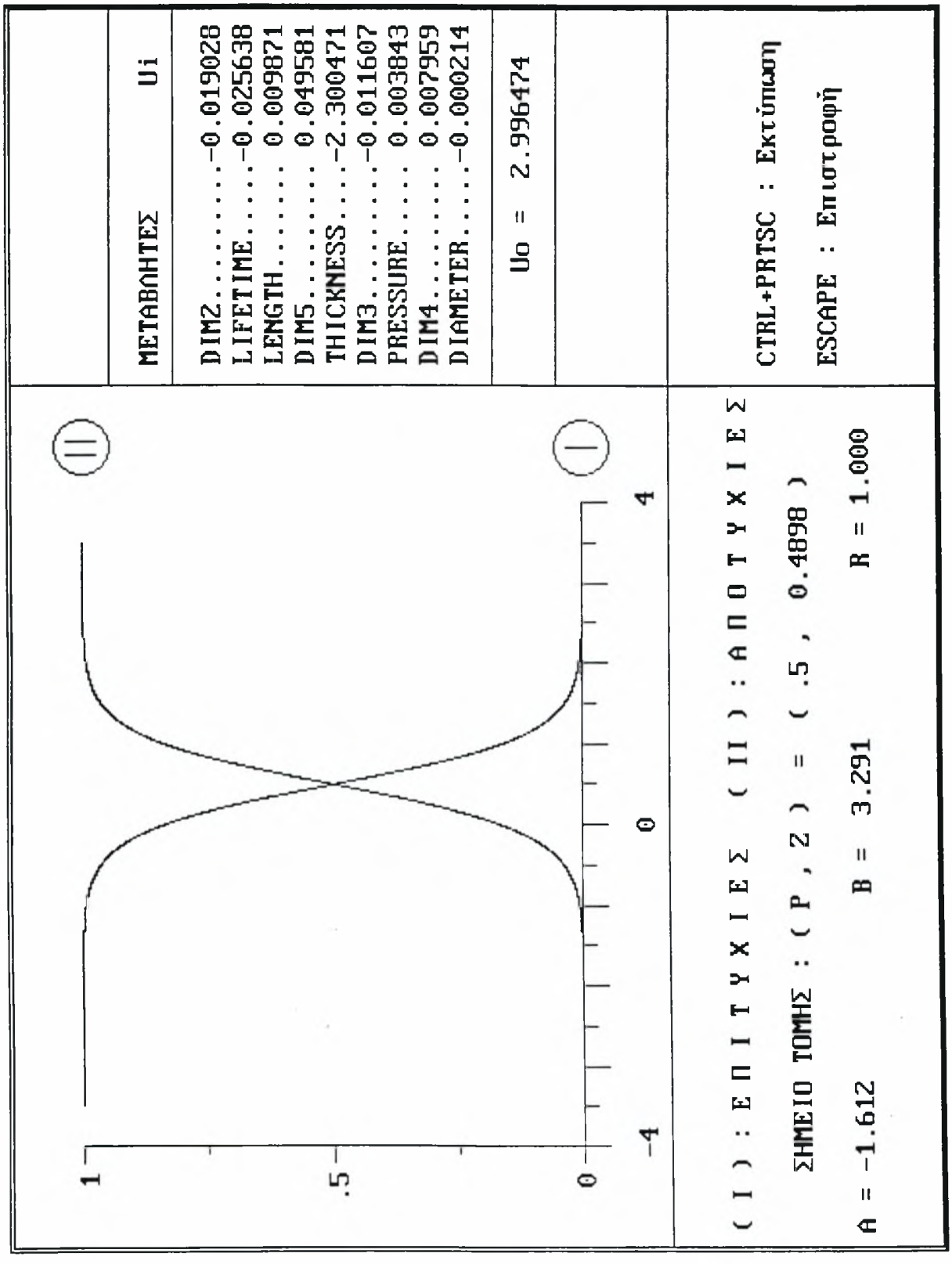
Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
ης Z

R A O ' S F			
Wilk's Λ	F	v1	v2
0.2857	43.8844	9	158

Πάτησε οποιοδήποτε πλήκτρο για επιστροφή

Πίνακας 8.4



Πίνακας 8.5

≡≡≡	Επιτυχίες	Αποτυχίες	Στατιστική Ανάλυση	Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

σων τιμών
ασπορών
ς συσχέτισης
τα της Z
ση

Ιδιοτιμές	C.C.C
0.00000	0.0000
2.49974	0.8451
-0.00000	0.0000
0.00000	0.0002
-0.00000	0.0000
-0.00000	0.0000
0.00000	0.0003
-0.00000	0.0000
-0.00000	0.0000

Πάτησε οποιοδήποτε πλήκτρο

Πίνακας 8.6

≡≡≡	Επιτυχίες	Αποτυχίες	Στατιστική Ανάλυση	Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
ης Z

B O X ' S F			
Box's B	F	v1	v2
1032.1985	21.4448	45	48524

Πάτησε οποιοδήποτε πλήκτρο για επιστροφή

Πίνακας 8.7

≡≡≡ Επιτυχίες Αποτυχίες Στατιστική Ανάλυση Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών

Ε λ ε γ χ ο ς	K o l m o g o r o v - S m i r n o v	
	Ε π ι τ υ χ ί ε ς	Α π ο τ υ χ ί ε ς
Μέση τιμή	-1.156271	2.136163
Μεταβλητότητα	1.141175	.7052858
Επ. σημαντικότητας	Ολα τα επίπεδα	Ολα τα επίπεδα

Πίνακας 8.8

≡≡≡ Επιτυχίες Αποτυχίες Στατιστική Ανάλυση Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
Κανονικότητα της Z
Παλινδρόμηση

Η εξίσωση παλινδρόμησης στις επιτυχίες έχει :
a = -1.6120 b = 3.2909 R = 1.0000
Αφού έγινε ο έλεγχος κατά Student προέκυψε
για τον πληθυσμό R > 0.

Πάτησε οποιοδήποτε πλήκτρο για συνέχεια.

Πίνακας 8.9

ΜΕΤΑΒΛΗΤΕΣ	ΜΕΣΗ ΤΙΜΗ	ΤΥΠ.ΑΠΟΚ/ΣΗ	ΜΕΓΙΣΤΟ	ΕΛΑΧΙΣΤΟ
LENGTH....	50.358	63.897	452.000	1.000
DIAMETER..	23.734	20.108	255.800	3.000
THICKNESS.	0.611	0.166	1.000	0.079
PRESSURE..	112.633	31.498	255.000	18.000
LIFETIME..	64.720	37.543	221.000	0.000
DIM1.....	41.773	24.566	230.000	14.353
DIM3.....	30.589	12.921	83.333	16.778
DIM4.....	13.574	7.812	34.668	0.484
DIM5.....	6.281	11.657	56.121	0.000

Πίνακας 9.1

ΜΕΤΑΒΛΗΤΕΣ	Ui	Si	ΚΑΤΑΤΑΞΗ	ΣΤΑΘΕΡΟΤΗΤΑ	Uo
LENGTH....	-0.000951	-0.060741	6	ΕΥΣΤΑΘΗΣ	
DIAMETER..	-0.001169	-0.023515	9	ΕΥΣΤΑΘΗΣ	
THICKNESS.	0.189385	0.031394	8	ΑΣΤΑΘΗΣ	
PRESSURE..	-0.003912	-0.123219	4	ΕΥΣΤΑΘΗΣ	
LIFETIME..	0.027480	1.031680	1	ΕΥΣΤΑΘΗΣ	0.046399
DIM1.....	-0.017134	-0.420916	3	ΑΣΤΑΘΗΣ	
DIM3.....	0.008658	0.111876	5	ΕΥΣΤΑΘΗΣ	
DIM4.....	-0.073751	-0.576134	2	ΕΥΣΤΑΘΗΣ	
DIM5.....	0.004403	0.051324	7	ΕΥΣΤΑΘΗΣ	

Πίνακας 9.2

Π Ι Ν Α Κ Α Σ Κ Α Ι Π Ο Σ Ο Σ Τ Α Τ Α Ξ Ι Ν Ο Μ Η Σ Η Σ

Ποσοστά Σωστής Ταξινόμησης

ΕΠΙΤΥΧΙΕΣ .... EF1 = 89.0%
ΑΠΟΤΥΧΙΕΣ .... EF2 = 94.9%
ΣΥΝΟΛΙΚΑ ..... EFt = 91.1%

		ΠΡΟΒΛΕΠΟΜΕΝΑ: ΕΠΙΤΥΧΙΕΣ	ΠΡΟΒΛΕΠΟΜΕΝΑ: ΑΠΟΤΥΧΙΕΣ	
ΠΡΑΓΜΑΤΙΚΑ	ΕΠΙΤΥΧΙΕΣ	97	12	109
	ΑΠΟΤΥΧΙΕΣ	3	56	59
		100	68	

ESCAPE : Επιστροφή

Πίνακας 9.3

≡≡≡ Επιτυχίες Αποτυχίες Στατιστική Ανάλυση Εξοδος

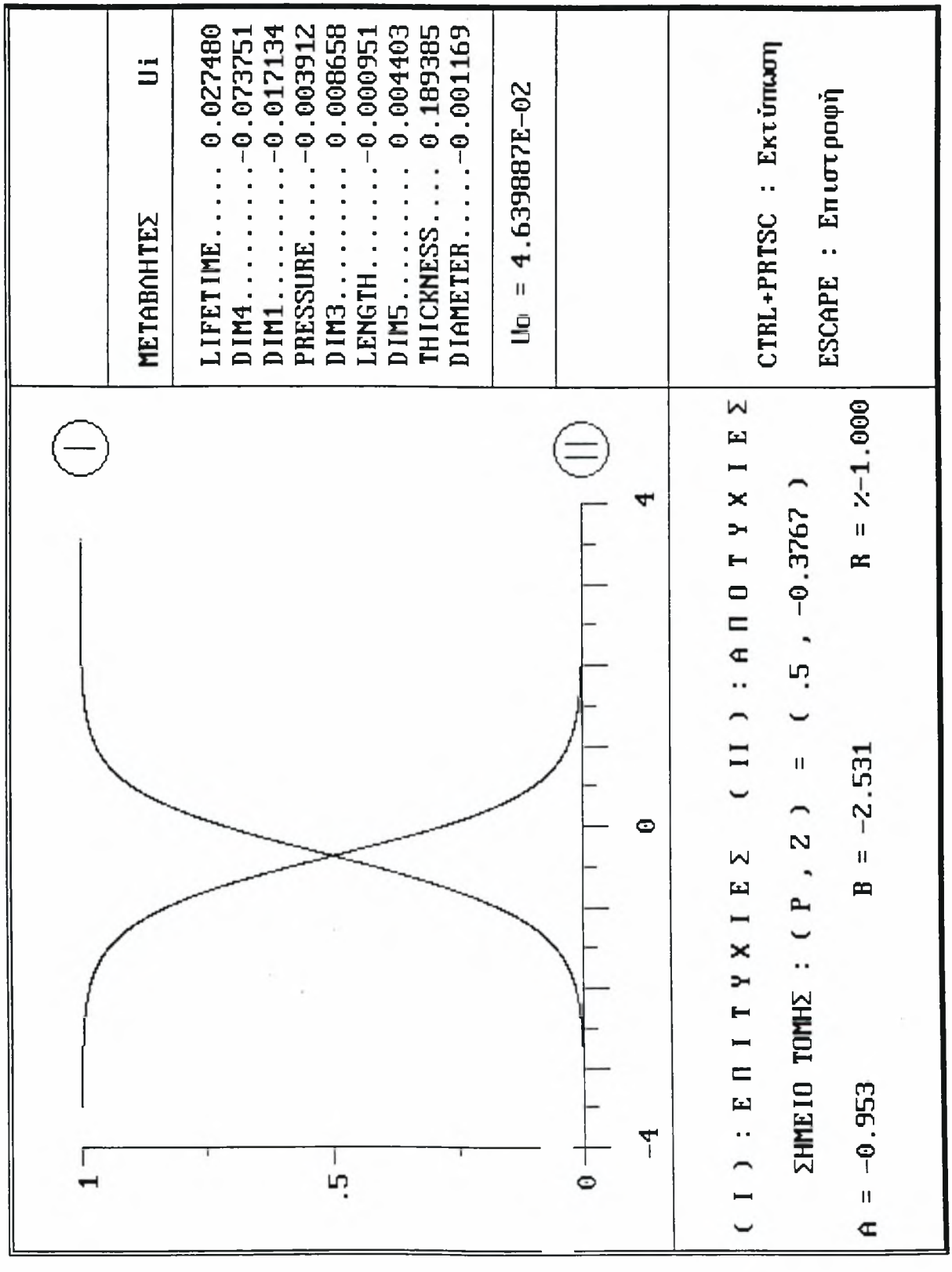
Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
ης Z

R A O ' S F			
Wilk's Λ	F	v1	v2
0.4036	25.9365	9	158

Πάτησε οποιοδήποτε πλήκτρο για επιστροφή

Πίνακας 9.4



Πίνακας 9.5

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
ης Z

B O X ' S F			
Box's B	F	v1	v2
683.2489	14.1951	45	48524

Πάτησε οποιοδήποτε πλήκτρο για επιστροφή

Πίνακας 9.6

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

σων τιμών
ασπορών
ς συσχέτισης
τα της Z
ση

Ιδιοτιμές	C.C.C
0.00000	0.0000
0.00000	0.0000
0.00000	0.0004
1.47740	0.7722
0.00000	0.0003
0.00000	0.0002
-0.00000	0.0000
-0.00000	0.0000
-0.00000	0.0000

Πάτησε οποιοδήποτε πλήκτρο

Πίνακας 9.7

≡≡≡ Επιτυχίες Αποτυχίες Στατιστική Ανάλυση Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών

Ε λ ε γ χ ο ς K o l m o g o r o v - S m i r n o v		
	Ε π ι τ υ χ ί ε ς	Α π ο τ υ χ ί ε ς
Μέση τιμή	.8889163	-1.642236
Μεταβλητότητα	1.146233	.6959417
Επ. σημαντικότητας	Ολα τα επίπεδα	Ολα τα επίπεδα

Πίνακας 9.8

≡≡≡ Επιτυχίες Αποτυχίες Στατιστική Ανάλυση Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
Κανονικότητα της Z
Παλινδρόμηση

Η εξίσωση παλινδρόμησης στις επιτυχίες έχει :
a = -0.9534 b = -2.5312 R = -0.9890
Αφού έγινε ο έλεγχος κατά Student προέκυψε
για τον πληθυσμό R = 0.

Πάτησε οποιοδήποτε πλήκτρο για συνέχεια.

Πίνακας 9.9

ΜΕΤΑΒΛΗΤΕΣ	ΜΕΣΗ ΤΙΜΗ	ΤΥΠ.ΑΠΟΚ/ΣΗ	ΜΕΓΙΣΤΟ	ΕΛΑΧΙΣΤΟ
LENGTH....	50.358	63.897	452.000	1.000
DIAMETER..	23.734	20.108	255.800	3.000
THICKNESS.	0.611	0.166	1.000	0.079
PRESSURE..	112.633	31.498	255.000	18.000
LIFETIME..	64.720	37.543	221.000	0.000
DIM1.....	41.773	24.566	230.000	14.353
DIM2.....	48.416	57.169	427.160	0.328
DIM4.....	13.574	7.812	34.668	0.484
DIM5.....	6.281	11.657	56.121	0.000

Πίνακας 10.1

ΜΕΤΑΒΛΗΤΕΣ	Ui	Si	ΚΑΤΑΤΑΞΗ	ΣΤΑΘΕΡΟΤΗΤΑ	Uo
LENGTH....	0.009353	0.597630	3	ΕΥΣΤΑΘΗΣ	
DIAMETER..	0.000874	0.017573	9	ΕΥΣΤΑΘΗΣ	
THICKNESS.	-2.298337	-0.380992	5	ΑΣΤΑΘΗΣ	
PRESSURE..	0.006602	0.207943	8	ΕΥΣΤΑΘΗΣ	
LIFETIME..	-0.026636	-0.999987	2	ΕΥΣΤΑΘΗΣ	1.559222
DIM1.....	0.012159	0.298682	6	ΑΣΤΑΘΗΣ	
DIM2.....	-0.018531	-1.059405	1	ΕΥΣΤΑΘΗΣ	
DIM4.....	0.031291	0.244437	7	ΕΥΣΤΑΘΗΣ	
DIM5.....	0.047407	0.552603	4	ΑΣΤΑΘΗΣ	

Πίνακας 10.2

Π Ι Ν Α Κ Α Σ Κ Α Ι Π Ο Σ Ο Σ Τ Α Τ Α Ξ Ι Ν Ο Μ Η Σ Η Σ

Ποσοστά Σωστής Ταξινόμησης

ΕΠΙΤΥΧΙΕΣ .... EF1 = 95.4%
ΑΠΟΤΥΧΙΕΣ .... EF2 = 98.3%
ΣΥΝΟΛΙΚΑ ..... EFt = 96.4%

		ΠΡΟΒΛΕΠΟΜΕΝΑ: ΕΠΙΤΥΧΙΕΣ	ΠΡΟΒΛΕΠΟΜΕΝΑ: ΑΠΟΤΥΧΙΕΣ	
ΠΡΑΓΜΑΤΙΚΑ	ΕΠΙΤΥΧΙΕΣ	104	5	109
	ΑΠΟΤΥΧΙΕΣ	1	58	59
		105	63	

ESCAPE : Επιστροφή

Πίνακας 10.3

≡≡≡ Επιτυχίες Αποτυχίες Στατιστική Ανάλυση Εξοδος

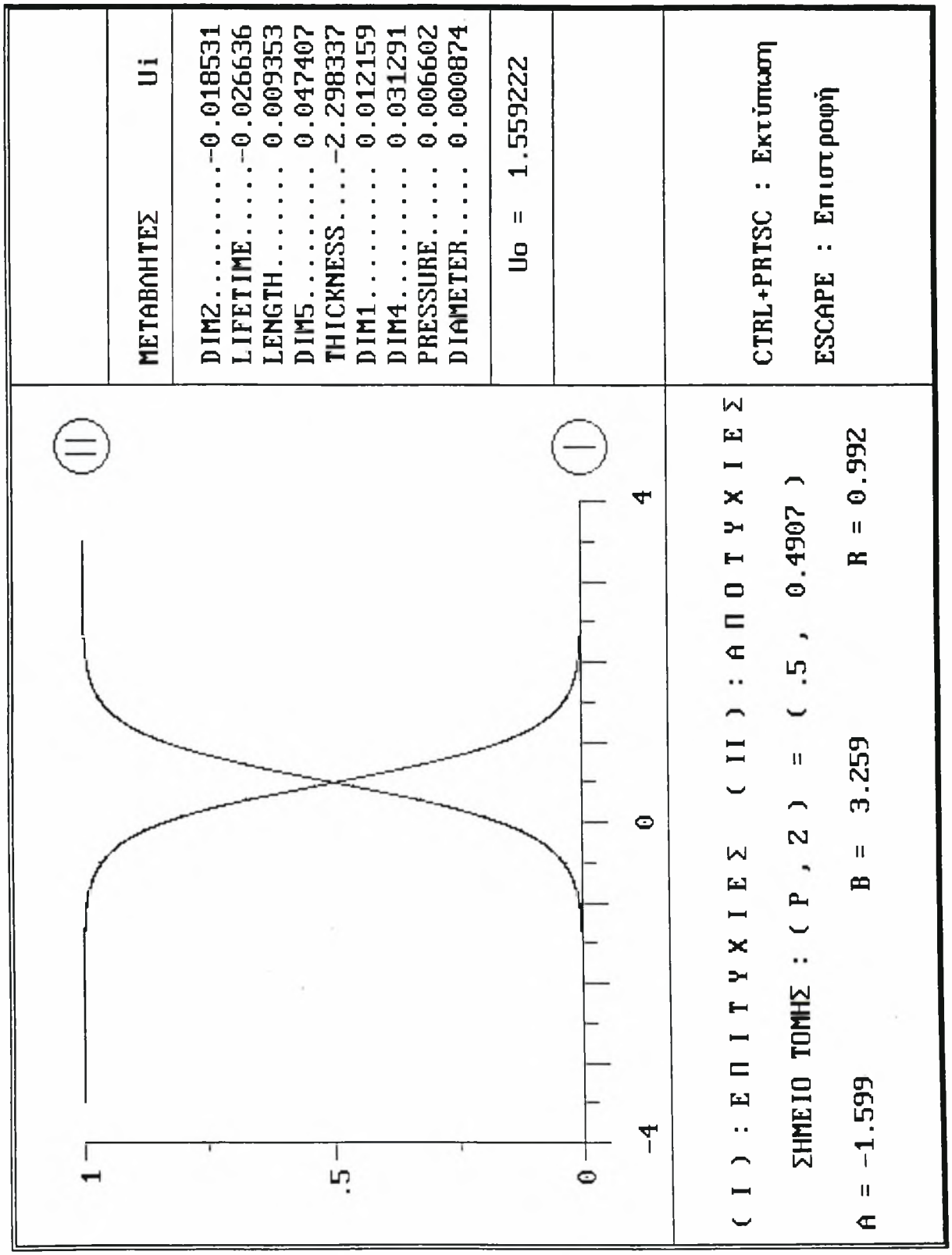
Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
ης Z

R A O ' S F			
Wilk's Λ	F	v1	v2
0.2796	45.2233	9	158

Πάτησε οποιοδήποτε πλήκτρο για επιστροφή

Πίνακας 10.4



Πίνακας 10.5

≡≡≡	Επιτυχίες	Αποτυχίες	Στατιστική Ανάλυση	Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
ης Z

B O X ' S F			
Box's B	F	v1	v2
998.1694	20.7378	45	48524
Πάτησε οποιοδήποτε πλήκτρο για επιστροφή			

Πίνακας 10.6

≡≡≡	Επιτυχίες	Αποτυχίες	Στατιστική Ανάλυση	Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

σων τιμών
ασπορών
ς συσχέτισης
τα της Z
ση

Ιδιοτιμές	C.C.C
0.00000	0.0000
2.57601	0.8487
-0.00000	0.0000
0.00000	0.0003
0.00000	0.0005
-0.00000	0.0000
0.00000	0.0002
-0.00000	0.0000
0.00000	0.0001
Πάτησε οποιοδήποτε πλήκτρο	

Πίνακας 10.7

≡≡≡	Επιτυχίες	Αποτυχίες	Στατιστική Ανάλυση	Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών

Ε λ ε γ χ ο ς	K o l m o g o r o v - S m i r n o v	
	Ε π ι τ υ χ ί ε ς	Α π ο τ υ χ ί ε ς
Μέση τιμή	-1.173778	2.168505
Μεταβλητότητα	1.117543	.7489455
Επ. σημαντικότητας	Ολα τα επίπεδα	Ολα τα επίπεδα

Πίνακας 10.8

≡≡≡	Επιτυχίες	Αποτυχίες	Στατιστική Ανάλυση	Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
Κανονικότητα της Z
Παλινδρόμηση

Η εξίσωση παλινδρόμησης στις επιτυχίες έχει :
a = -1.5995 b = 3.2595 R = 0.9920
Αφού έγινε ο έλεγχος κατά Student προέκυψε
για τον πληθυσμό R > 0.

Πάτησε οποιοδήποτε πλήκτρο για συνέχεια.

Πίνακας 10.9

ΜΕΤΑΒΛΗΤΕΣ	ΜΕΣΗ ΤΙΜΗ	ΤΥΠ.ΑΠΟΚ/ΣΗ	ΜΕΓΙΣΤΟ	ΕΛΑΧΙΣΤΟ
LENGTH....	50.358	63.897	452.000	1.000
DIAMETER..	23.734	20.108	255.800	3.000
THICKNESS.	0.611	0.166	1.000	0.079
PRESSURE..	112.633	31.498	255.000	18.000
LIFETIME..	64.720	37.543	221.000	0.000
DIM1.....	41.773	24.566	230.000	14.353
DIM2.....	48.416	57.169	427.160	0.328
DIM3.....	30.589	12.921	83.333	16.778
DIM5.....	6.281	11.657	56.121	0.000

Πίνακας 11.1

ΜΕΤΑΒΛΗΤΕΣ	Ui	Si	ΚΑΤΑΤΑΞΗ	ΣΤΑΘΕΡΟΤΗΤΑ	Uo
LENGTH....	0.009943	0.635329	3	ΕΥΣΤΑΘΗΣ	
DIAMETER..	0.000148	0.002975	9	ΕΥΣΤΑΘΗΣ	
THICKNESS.	-2.379670	-0.394475	5	ΑΣΤΑΘΗΣ	
PRESSURE..	0.010964	0.345335	6	ΕΥΣΤΑΘΗΣ	
LIFETIME..	-0.026104	-0.980027	2	ΕΥΣΤΑΘΗΣ	2.040960
DIM1.....	0.010352	0.254314	7	ΑΣΤΑΘΗΣ	
DIM2.....	-0.019057	-1.089468	1	ΕΥΣΤΑΘΗΣ	
DIM3.....	-0.014902	-0.192551	8	ΕΥΣΤΑΘΗΣ	
DIM5.....	0.049198	0.573486	4	ΑΣΤΑΘΗΣ	

Πίνακας 11.2

Π Ι Ν Α Κ Α Σ Κ Α Ι Π Ο Σ Ο Σ Τ Α Τ Α Ξ Ι Ν Ο Μ Η Σ Η Σ

Ποσοστά Σωστής Ταξινόμησης

ΕΠΙΤΥΧΙΕΣ .... EF1 = 96.3%
ΑΠΟΤΥΧΙΕΣ .... EF2 = 94.9%
ΣΥΝΟΛΙΚΑ ..... EFt = 95.8%

		ΠΡΟΒΛΕΠΟΜΕΝΑ: ΕΠΙΤΥΧΙΕΣ	ΠΡΟΒΛΕΠΟΜΕΝΑ: ΑΠΟΤΥΧΙΕΣ	
ΠΡΑΓΜΑΤΙΚΑ	ΕΠΙΤΥΧΙΕΣ	105	4	109
	ΑΠΟΤΥΧΙΕΣ	3	56	59
		108	60	

ESCAPE : Επιστροφή

Πίνακας 11.3

=== Επιτυχίες Αποτυχίες Στατιστική Ανάλυση Εξοδος

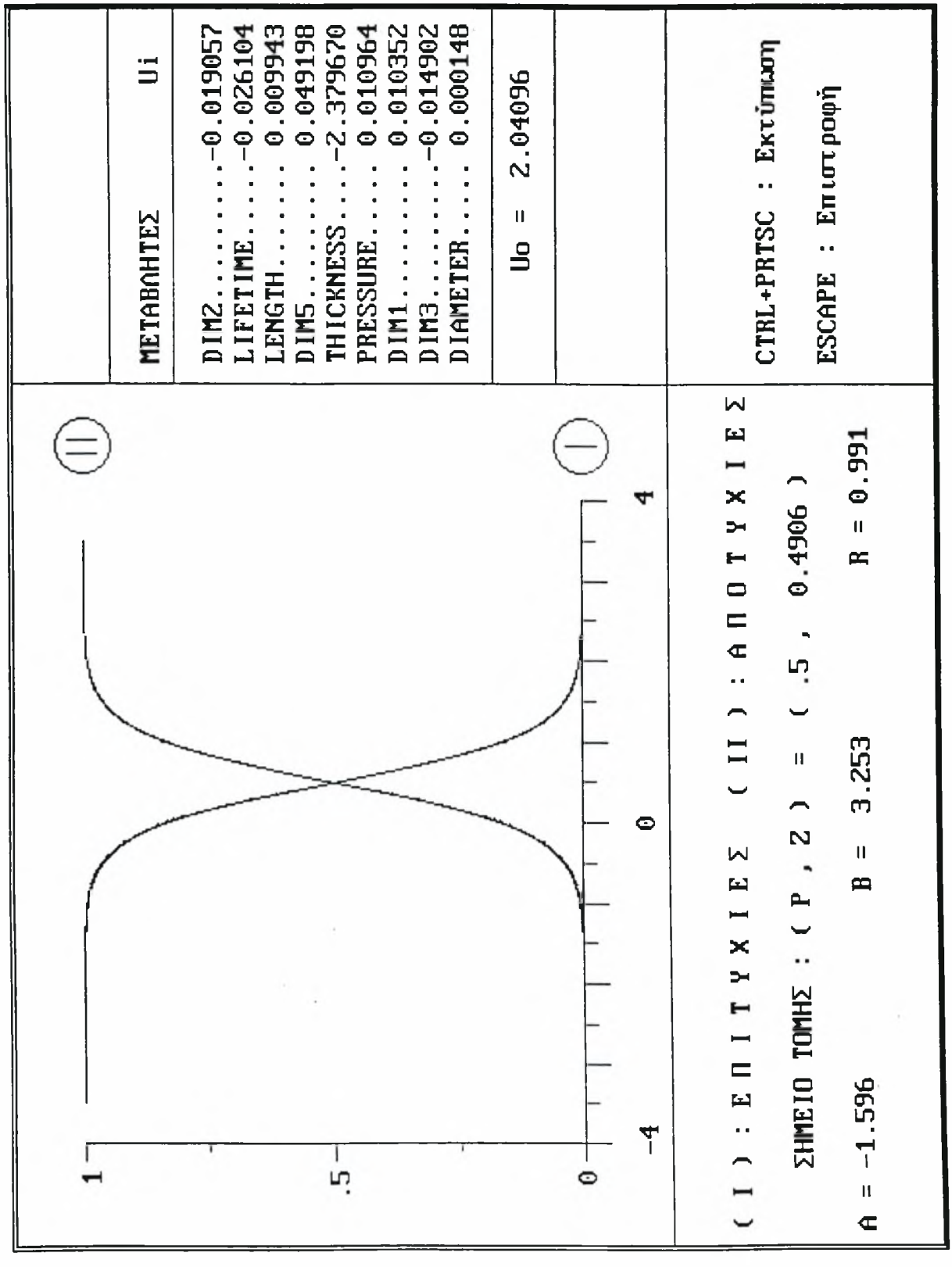
Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
ης Z

R A O ' S F			
Wilk's Λ	F	v1	v2
0.2800	45.1460	9	158

Πάτησε οποιοδήποτε πλήκτρο για επιστροφή

Πίνακας 11.4



Πίνακας 11.5

≡≡≡	Επιτυχίες	Αποτυχίες	Στατιστική Ανάλυση	Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
ης Z

B O X ' S F			
Box's B	F	v1	v2
1120.3269	23.2757	45	48524
Πάτησε οποιοδήποτε πλήκτρο για επιστροφή			

Πίνακας 11.6

≡≡≡	Επιτυχίες	Αποτυχίες	Στατιστική Ανάλυση	Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

σων τιμών
ασπορών
ς συσχέτισης
τα της Z
ση

Ιδιοτιμές	C.C.C
-59.07001	0.0000
2.57161	0.8485
-0.00000	0.0000
0.00000	0.0003
0.00000	0.0004
-0.00000	0.0000
0.00000	0.0000
0.00000	0.0000
-0.00000	0.0000
Πάτησε οποιοδήποτε πλήκτρο	

Πίνακας 11.7

```
===    Επιτυχίες    Αποτυχίες    Στατιστική Ανάλυση    Εξοδος
                                  | Συνολικά στατιστικά στοιχεία
                                  | Διαχωριστική ανάλυση
                                  | Πίνακας ταξινόμησης
                                  | ξινόμησης
                                  | κασίας
                 Ισότητα μέσων τιμών
                 Ισότητα διασπορών
```

Ε λ ε γ χ ο ς	K o l m o g o r o v - S m i r n o v	
	Ε π ι τ υ χ ί ε ς	Α π ο τ υ χ ί ε ς
Μέση τιμή	-1.172774	2.166652
Μεταβλητότητα	1.124114	.7368065
Επ. σημαντικότητας	Ολα τα επίπεδα	Ολα τα επίπεδα

Πίνακας 11.8

```
===    Επιτυχίες    Αποτυχίες    Στατιστική Ανάλυση    Εξοδο
                                  | Συνολικά στατιστικά στοιχεία
                                  | Διαχωριστική ανάλυση
                                  | Πίνακας ταξινόμησης
                                  | ξινόμησης
                                  | κασίας
                 Ισότητα μέσων τιμών
                 Ισότητα διασπορών
                 Συντελεστης συσχέτισης
                 Κανονικότητα της Z
                 Παλινδρόμηση

Η εξίσωση παλινδρόμησης στις επιτυχίες έχει :
 a = -1.5961     b =  3.2531     R =  0.9913
 Αφού έγινε ο έλεγχος κατά Student προέκυψε
 για τον πληθυσμό R > 0.

   Πάτησε οποιοδήποτε πλήκτρο για συνέχεια.
```

Πίνακας 11.9

ΜΕΤΑΒΛΗΤΕΣ	ΜΕΣΗ ΤΙΜΗ	ΤΥΠ.ΑΠΟΚ/ΣΗ	ΜΕΓΙΣΤΟ	ΕΛΑΧΙΣΤΟ
LENGTH....	50.358	63.897	452.000	1.000
DIAMETER..	23.734	20.108	255.800	3.000
THICKNESS.	0.611	0.166	1.000	0.079
PRESSURE..	112.633	31.498	255.000	18.000
LIFETIME..	64.720	37.543	221.000	0.000
DIM1.....	41.773	24.566	230.000	14.353
DIM2.....	48.416	57.169	427.160	0.328
DIM3.....	30.589	12.921	83.333	16.778
DIM4.....	13.574	7.812	34.668	0.484

Πίνακας 12.1

ΜΕΤΑΒΛΗΤΕΣ	Ui	Si	ΚΑΤΑΤΑΞΗ	ΣΤΑΘΕΡΟΤΗΤΑ	Uo
LENGTH....	0.010679	0.682376	3	ΕΥΣΤΑΘΗΣ	
DIAMETER..	0.000534	0.010738	9	ΕΥΣΤΑΘΗΣ	
THICKNESS.	-1.690891	-0.280297	6	ΑΣΤΑΘΗΣ	
PRESSURE..	0.009521	0.299907	5	ΕΥΣΤΑΘΗΣ	
LIFETIME..	-0.021894	-0.821963	2	ΕΥΣΤΑΘΗΣ	0.793714
DIM1.....	0.013459	0.330631	4	ΑΣΤΑΘΗΣ	
DIM2.....	-0.014707	-0.840755	1	ΕΥΣΤΑΘΗΣ	
DIM3.....	-0.009755	-0.126047	8	ΕΥΣΤΑΘΗΣ	
DIM4.....	0.035466	0.277054	7	ΕΥΣΤΑΘΗΣ	

Πίνακας 12.2

Π Ι Ν Α Κ Α Σ Κ Α Ι Π Ο Σ Ο Σ Τ Α Τ Α Ξ Ι Ν Ο Μ Η Σ Η Σ

Ποσοστά Σωστής Ταξινόμησης

ΕΠΙΤΥΧΙΕΣ .... EF1 = 94.5%
ΑΠΟΤΥΧΙΕΣ .... EF2 = 96.6%
ΣΥΝΟΛΙΚΑ ..... EFt = 95.2%

		ΠΡΟΒΛΕΠΟΜΕΝΑ: ΕΠΙΤΥΧΙΕΣ	ΠΡΟΒΛΕΠΟΜΕΝΑ: ΑΠΟΤΥΧΙΕΣ	
ΠΡΑΓΜΑΤΙΚΑ	ΕΠΙΤΥΧΙΕΣ	103	6	109
	ΑΠΟΤΥΧΙΕΣ	2	57	59
		105	63	

ESCAPE : Επιστροφή

Πίνακας 12.3

≡≡≡ Επιτυχίες Αποτυχίες Στατιστική Ανάλυση Εξοδος

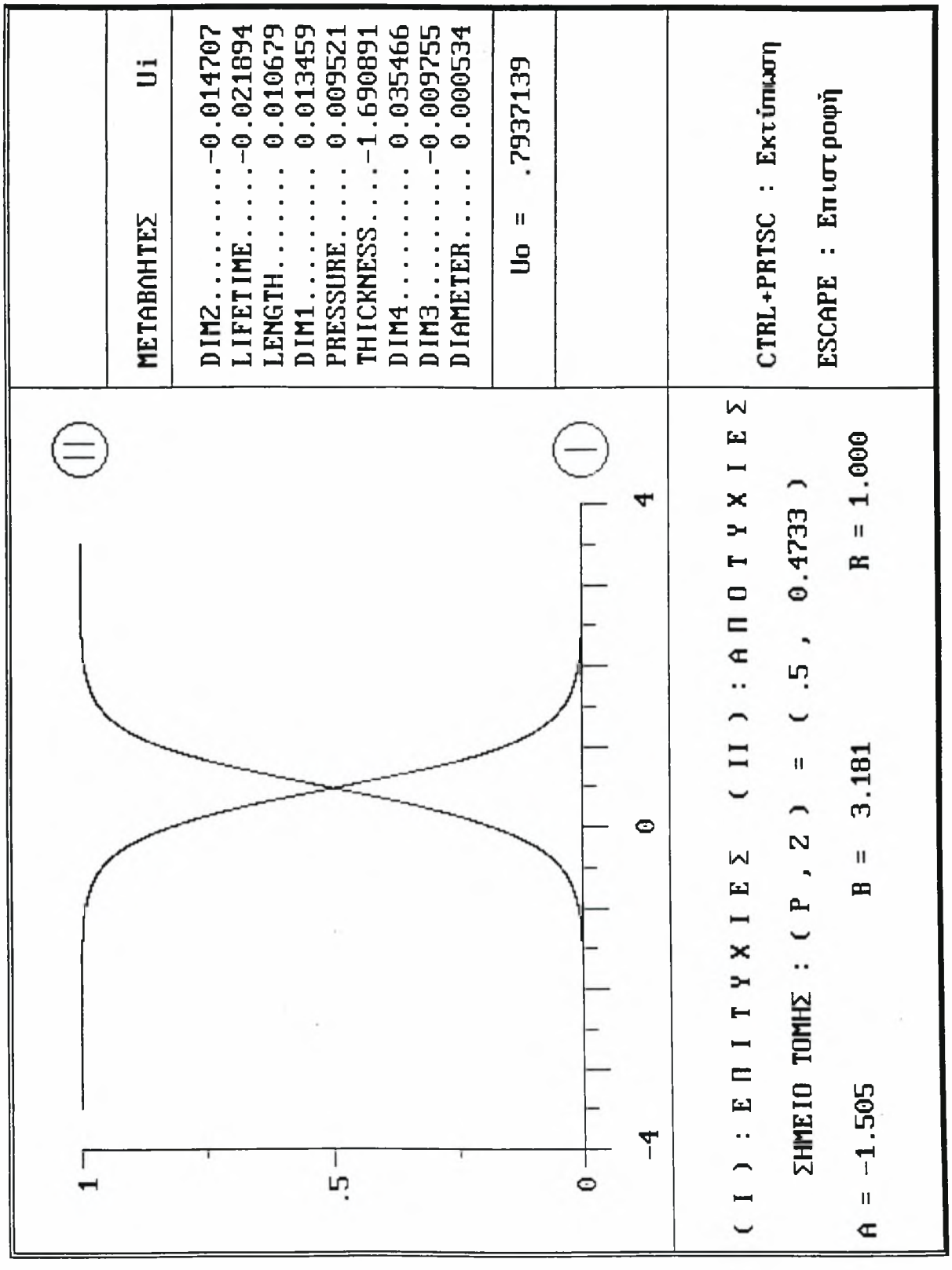
Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
ης Z

R A O ' S F			
Wilk's Λ	F	v1	v2
0.3002	40.9330	9	158

Πάτησε οποιοδήποτε πλήκτρο για επιστροφή

Πίνακας 12.4



Πίνακας 12.5

≡≡≡	Επιτυχίες	Αποτυχίες	Στατιστική Ανάλυση	Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
ης Z

B O X ' S F			
Box's B	F	v1	v2
999.7413	20.7705	45	48524
Πάτησε οποιοδήποτε πλήκτρο για επιστροφή			

Πίνακας 12.6

≡≡≡	Επιτυχίες	Αποτυχίες	Στατιστική Ανάλυση	Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

σων τιμών
ασπορών
ς συσχέτισης
τα της Z
ση

Ιδιοτιμές	C.C.C
0.00000	0.0000
2.33163	0.8366
0.00000	0.0003
-0.00000	0.0000
0.00000	0.0004
0.00000	0.0001
-0.00000	0.0000
-0.00000	0.0000
0.00000	0.0001
Πάτησε οποιοδήποτε πλήκτρο	

Πίνακας 12.7

≡≡≡	Επιτυχίες	Αποτυχίες	Στατιστική Ανάλυση	Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών

Ε λ ε γ χ ο ς	K o l m o g o r o v - S m i r n o v	
	Ε π ι τ υ χ ί ε ς	Α π ο τ υ χ ί ε ς
Μέση τιμή	-1.116713	2.063081
Μεταβλητότητα	.9965853	.97241
Επ. σημαντικότητας	Ολα τα επίπεδα	Ολα τα επίπεδα

Πίνακας 12.8

≡≡≡	Επιτυχίες	Αποτυχίες	Στατιστική Ανάλυση	Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
Κανονικότητα της Z
Παλινδρόμηση

Η εξίσωση παλινδρόμησης στις επιτυχίες έχει :
a = -1.5054 b = 3.1807 R = 1.0000
Αφού έγινε ο έλεγχος κατά Student προέκυψε
για τον πληθυσμό R > 0.

Πάτησε οποιοδήποτε πλήκτρο για συνέχεια.

Πίνακας 12.9

ΜΕΤΑΒΛΗΤΕΣ	ΜΕΣΗ ΤΙΜΗ	ΤΥΠ.ΑΠΟΚ/ΣΗ	ΜΕΓΙΣΤΟ	ΕΛΑΧΙΣΤΟ
DIM1.....	41.773	24.566	230.000	14.353
DIM2.....	48.416	57.169	427.160	0.328
DIM3.....	30.589	12.921	83.333	16.778
DIM4.....	13.574	7.812	34.668	0.484
DIM5.....	6.281	11.657	56.121	0.000

Πίνακας 13.1

```
===   Επιτυχίες   Αποτυχίες   Στατιστική Ανάλυση   Εξοδος

                              | Συνολικά στατιστικά στοιχεία
                              | Διαχωριστική ανάλυση
                              | Πίνακας ταξινόμησης
                                ξινόμησης
                                κασίας
                 | Ισότητα μέσων τιμών
                 | Ισότητα διασπορών
                                ης
 | Διακόπτεται η μαθηματική διαδικασία επειδή
 | δεν υπάρχει ο αντίστροφος του πίνακα W.
 | Πάτησε οποιοδήποτε πλήκτρο
```

Πίνακας 13.2

ΜΕΤΑΒΛΗΤΕΣ	ΜΕΣΗ ΤΙΜΗ	ΤΥΠ.ΑΠΟΚ/ΣΗ	ΜΕΓΙΣΤΟ	ΕΛΑΧΙΣΤΟ
LENGTH....	50.358	63.897	452.000	1.000
THICKNESS.	0.611	0.166	1.000	0.079
LIFETIME..	64.720	37.543	221.000	0.000
DIM2.....	48.416	57.169	427.160	0.328
DIM5.....	6.281	11.657	56.121	0.000

Πίνακας 14.1

ΜΕΤΑΒΛΗΤΕΣ	Ui	Si	ΚΑΤΑΤΑΞΗ	ΣΤΑΘΕΡΟΤΗΤΑ	Uo
LENGTH....	0.011107	0.709673	3	ΕΥΣΤΑΘΗΣ	3.237785
THICKNESS.	-2.422224	-0.401529	5	ΑΣΤΑΘΗΣ	
LIFETIME..	-0.026377	-0.990278	2	ΕΥΣΤΑΘΗΣ	
DIM2.....	-0.020203	-1.154991	1	ΕΥΣΤΑΘΗΣ	
DIM5.....	0.058546	0.682458	4	ΑΣΤΑΘΗΣ	

Πίνακας 14.2

Π Ι Ν Α Κ Α Σ Κ Α Ι Π Ο Σ Ο Σ Τ Α Τ Α Ξ Ι Ν Ο Μ Η Σ Η Σ

Ποσοστά Σωστής Ταξινόμησης

ΕΠΙΤΥΧΙΕΣ .... EF1 = 96.3%
ΑΠΟΤΥΧΙΕΣ .... EF2 = 94.9%
ΣΥΝΟΛΙΚΑ ..... EFt = 95.8%

		ΠΡΟΒΛΕΠΟΜΕΝΑ: ΕΠΙΤΥΧΙΕΣ	ΠΡΟΒΛΕΠΟΜΕΝΑ: ΑΠΟΤΥΧΙΕΣ	
ΠΡΑΓΜΑΤΙΚΑ	ΕΠΙΤΥΧΙΕΣ	105	4	109
	ΑΠΟΤΥΧΙΕΣ	3	56	59
		108	60	

ESCAPE : Επιστροφή

Πίνακας 14.3

≡≡≡ Επιτυχίες Αποτυχίες Στατιστική Ανάλυση Εξοδος

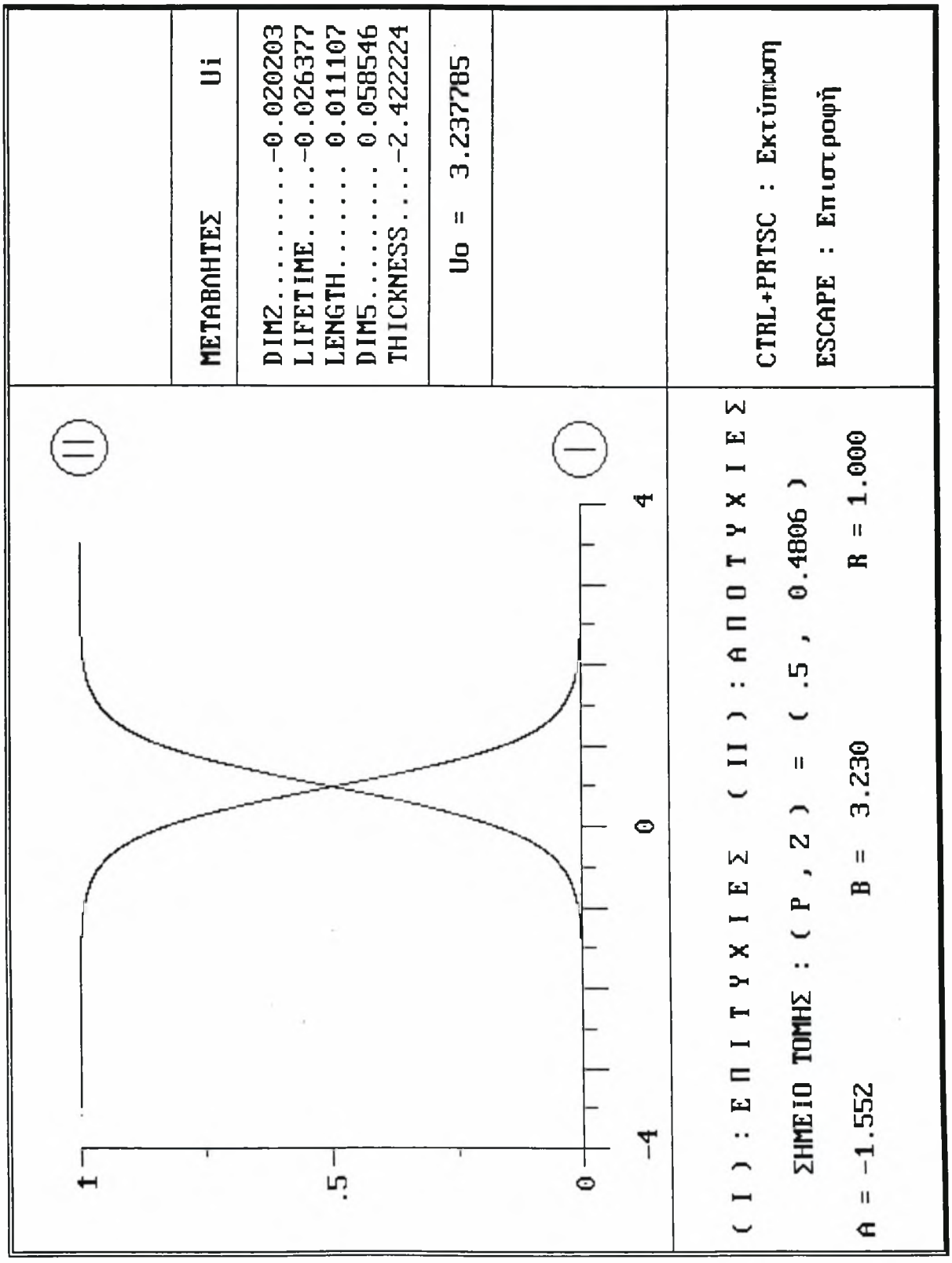
Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
ης Z

R A O ' S F			
Wilk's Λ	F	v1	v2
0.2937	77.9246	5	162

Πάτησε οποιοδήποτε πλήκτρο για επιστροφή

Πίνακας 14.4



Πίνακας 14.5

≡≡≡ Επιτυχίες Αποτυχίες Στατιστική Ανάλυση Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
ης Z

B O X ' S F			
Box's B	F	v1	v2
650.5723	41.7814	15	58057
Πάτησε οποιοδήποτε πλήκτρο για επιστροφή			

Πίνακας 14.6

≡≡≡ Επιτυχίες Αποτυχίες Στατιστική Ανάλυση Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

σων τιμών
ασπορών
ς συσχέτισης
τα της Z
ση

Ιδιοτιμές	C.C.C
-58.20872	0.0000
2.40508	0.8404
0.00000	0.0003
0.00000	0.0000
0.00000	0.0000
Πάτησε οποιοδήποτε πλήκτρο	

Πίνακας 14.7

≡≡≡	Επιτυχίες	Αποτυχίες	Στατιστική Ανάλυση	Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών

Ε λ ε γ χ ο ς	K o l m o g o r o v - S m i r n o v	
	Ε π ι τ υ χ ί ε ς	Α π ο τ υ χ ί ε ς
Μέση τιμή	-1.134167	2.095325
Μεταβλητότητα	1.149971	.6890364
Επ. σημαντικότητας	Ολα τα επίπεδα	Ολα τα επίπεδα

Πίνακας 14.8

≡≡≡	Επιτυχίες	Αποτυχίες	Στατιστική Ανάλυση	Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
Κανονικότητα της Z
Παλινδρόμηση

Η εξίσωση παλινδρόμησης στις επιτυχίες έχει :
a = -1.5523 b = 3.2299 R = 1.0000
Αφού έγινε ο έλεγχος κατά Student προέκυψε
για τον πληθυσμό R > 0.

Πάτησε οποιοδήποτε πλήκτρο για συνέχεια.

Πίνακας 14.9

ΜΕΤΑΒΛΗΤΕΣ	ΜΕΣΗ ΤΙΜΗ	ΤΥΠ.ΑΠΟΚ/ΣΗ	ΜΕΓΙΣΤΟ	ΕΛΑΧΙΣΤΟ
DIAMETER..	23.734	20.108	255.800	3.000
PRESSURE..	112.633	31.498	255.000	18.000
DIM1.....	41.773	24.566	230.000	14.353
DIM3.....	30.589	12.921	83.333	16.778
DIM4.....	13.574	7.812	34.668	0.484

Πίνακας 15.1

ΜΕΤΑΒΛΗΤΕΣ	Ui	Si	ΚΑΤΑΤΑΞΗ	ΣΤΑΘΕΡΟΤΗΤΑ	Uo
DIAMETER..	-0.009642	-0.193873	4	ΑΣΤΑΘΗΣ	
PRESSURE..	0.025433	0.801097	1	ΕΥΣΤΑΘΗΣ	
DIM1.....	0.006007	0.147566	5	ΑΣΤΑΘΗΣ	-2.068495
DIM3.....	-0.042633	-0.550862	2	ΕΥΣΤΑΘΗΣ	
DIM4.....	0.035798	0.279646	3	ΕΥΣΤΑΘΗΣ	

Πίνακας 15.2

Π Ι Ν Α Κ Α Σ Κ Α Ι Π Ο Σ Ο Σ Τ Α Τ Α Ξ Ι Ν Ο Μ Η Σ Η Σ

Ποσοστά Σωστής Ταξινόμησης

ΕΠΙΤΥΧΙΕΣ .... EF1 = 65.1%
ΑΠΟΤΥΧΙΕΣ .... EF2 = 84.7%
ΣΥΝΟΛΙΚΑ ..... EFt = 72.0%

		ΠΡΟΒΛΕΠΟΜΕΝΑ: ΕΠΙΤΥΧΙΕΣ	ΠΡΟΒΛΕΠΟΜΕΝΑ: ΑΠΟΤΥΧΙΕΣ	
ΠΡΑΓΜΑΤΙΚΑ	ΕΠΙΤΥΧΙΕΣ	71	38	109
	ΑΠΟΤΥΧΙΕΣ	9	50	59
		80	88	

ESCAPE : Επιστροφή

Πίνακας 15.3

≡≡≡ Επιτυχίες Αποτυχίες Στατιστική Ανάλυση Εξοδος

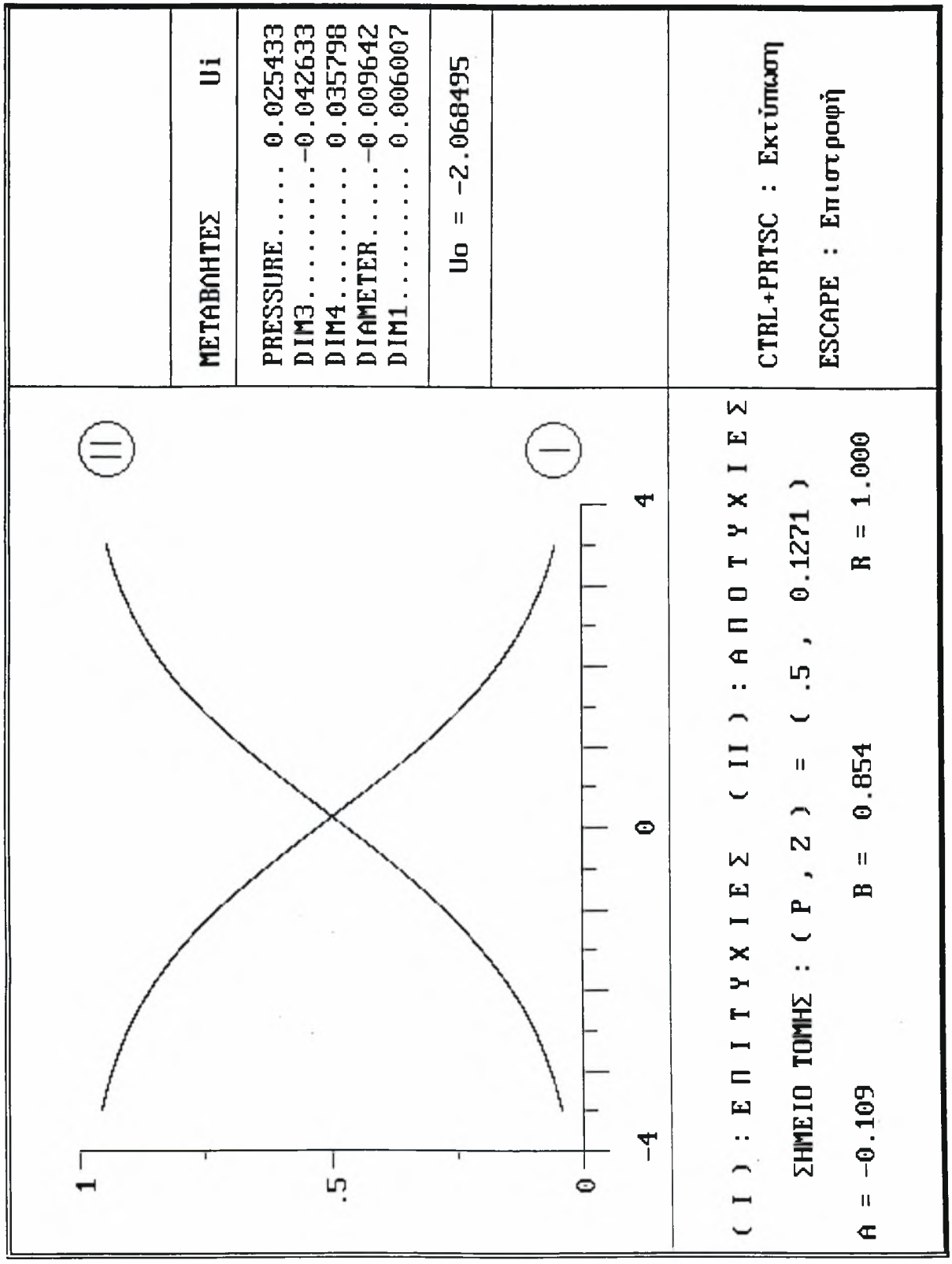
Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
ης Z

R A O ' S F			
Wilk's Λ	F	v1	v2
0.8561	5.4478	5	162

Πάτησε οποιοδήποτε πλήκτρο για επιστροφή

Πίνακας 15.4



Πίνακας 15.5

≡≡≡ Επιτυχίες Αποτυχίες Στατιστική Ανάλυση Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
ης Z

B O X ' S F			
Box's B	F	v1	v2
234.4877	15.0594	15	58057
Πάτησε οποιοδήποτε πλήκτρο για επιστροφή			

Πίνακας 15.6

≡≡≡ Επιτυχίες Αποτυχίες Στατιστική Ανάλυση Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

σων τιμών
ασπορών
ς συσχέτισης
τα της Z
ση

Ιδιοτιμές	C.C.C
0.00000	0.0000
0.16814	0.3794
0.00000	0.0004
-0.00000	0.0000
0.00000	0.0002
Πάτησε οποιοδήποτε πλήκτρο	

Πίνακας 15.7

≡≡≡	Επιτυχίες	Αποτυχίες	Στατιστική Ανάλυση	Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών

Ε λ ε γ χ ο ς	K o l m o g o r o v - S m i r n o v	
	Ε π ι τ υ χ ί ε ς	Α π ο τ υ χ ί ε ς
Μέση τιμή	-.2998808	.5540169
Μεταβλητότητα	1.149741	.6894604
Επ. σημαντικότητας	Ολα τα επίπεδα	Ολα τα επίπεδα

Πίνακας 15.8

≡≡≡	Επιτυχίες	Αποτυχίες	Στατιστική Ανάλυση	Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
Κανονικότητα της Z
Παλινδρόμηση

Η εξίσωση παλινδρόμησης στις επιτυχίες έχει :
a = -0.1085 b = 0.8539 R = 1.0000
Αφού έγινε ο έλεγχος κατά Student προέκυψε
για τον πληθυσμό R > 0.

Πάτησε οποιοδήποτε πλήκτρο για συνέχεια.

Πίνακας 15.9

ΜΕΤΑΒΛΗΤΕΣ	ΜΕΣΗ ΤΙΜΗ	ΤΥΠ.ΑΠΟΚ/ΣΗ	ΜΕΓΙΣΤΟ	ΕΛΑΧΙΣΤΟ
LENGTH....	50.358	63.897	452.000	1.000
DIM1.....	41.773	24.566	230.000	14.353
DIM2.....	48.416	57.169	427.160	0.328
DIM3.....	30.589	12.921	83.333	16.778
DIM4.....	13.574	7.812	34.668	0.484
DIM5.....	6.281	11.657	56.121	0.000

Πίνακας 16.1

ΜΕΤΑΒΛΗΤΕΣ	Ui	Si	ΚΑΤΑΤΑΞΗ	ΣΤΑΘΕΡΟΤΗΤΑ	Uo
LENGTH....	-0.014886	-0.951150	2	ΕΥΣΤΑΘΗΣ	0.041677
DIM1.....	0.004623	0.113563	5	ΕΥΣΤΑΘΗΣ	
DIM2.....	0.017202	0.983420	1	ΕΥΣΤΑΘΗΣ	
DIM3.....	0.000365	0.004717	6	ΕΥΣΤΑΘΗΣ	
DIM4.....	-0.039152	-0.305848	4	ΕΥΣΤΑΘΗΣ	
DIM5.....	0.032199	0.375340	3	ΕΥΣΤΑΘΗΣ	

Πίνακας 16.2

Π Ι Ν Α Κ Α Σ Κ Α Ι Π Ο Σ Ο Σ Τ Α Τ Α Ξ Ι Ν Ο Μ Η Σ Η Σ

Ποσοστά Σωστής Ταξινόμησης

ΕΠΙΤΥΧΙΕΣ .... EF1 = 90.8%
ΑΠΟΤΥΧΙΕΣ .... EF2 = 86.4%
ΣΥΝΟΛΙΚΑ ..... EFt = 89.3%

		ΠΡΟΒΛΕΠΟΜΕΝΑ: ΕΠΙΤΥΧΙΕΣ	ΠΡΟΒΛΕΠΟΜΕΝΑ: ΑΠΟΤΥΧΙΕΣ	
ΠΡΑΓΜΑΤΙΚΑ	ΕΠΙΤΥΧΙΕΣ	99	10	109
	ΑΠΟΤΥΧΙΕΣ	8	51	59
		107	61	

ESCAPE : Επιστροφή

Πίνακας 16.3

≡≡≡ Επιτυχίες Αποτυχίες Στατιστική Ανάλυση Εξοδος

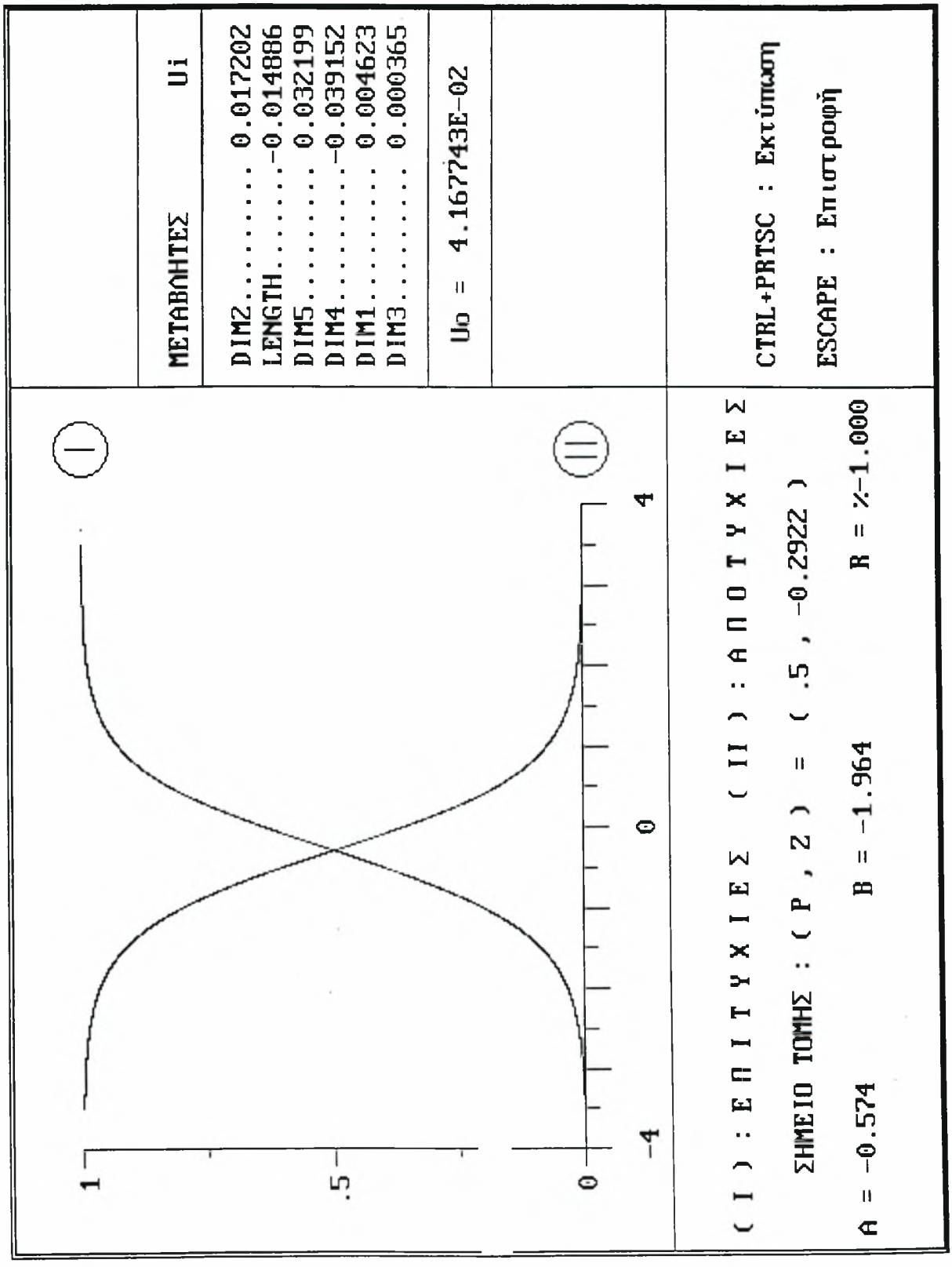
Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
ης Z

R A O ' S F			
Wilk's Λ	F	v1	v2
0.5293	23.8655	6	161

Πάτησε οποιοδήποτε πλήκτρο για επιστροφή

Πίνακας 16.4



Πίνακας 16.5

≡≡≡	Επιτυχίες	Αποτυχίες	Στατιστική Ανάλυση	Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
ης Z

B O X ' S F			
Box's B	F	v1	v2
694.0638	31.6036	21	53659
Πάτησε οποιοδήποτε πλήκτρο για επιστροφή			

Πίνακας 16.6

≡≡≡	Επιτυχίες	Αποτυχίες	Στατιστική Ανάλυση	Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

σων τιμών
ασπορών
ς συσχέτισης
τα της Z
ση

Ιδιοτιμές	C.C.C
0.00000	0.0000
0.00000	0.0000
-0.00000	0.0000
0.88940	0.6861
0.00000	0.0002
-0.00000	0.0000
Πάτησε οποιοδήποτε πλήκτρο	

Πίνακας 16.7

```
===    Επιτυχίες    Αποτυχίες    Στατιστική Ανάλυση    Εξοδος
                                  Συνολικά στατιστικά στοιχεία
                                  Διαχωριστική ανάλυση
                                  Πίνακας ταξινόμησης
                                  ξινόμησης
                                  κασίας
                Ισότητα μέσων τιμών
                Ισότητα διασπορών
```

Ε λ ε γ χ ο ς	K o l m o g o r o v - S m i r n o v	
	Ε π ι τ υ χ ί ε ς	Α π ο τ υ χ ί ε ς
Μέση τιμή	.6897	-1.274192
Μεταβλητότητα	.7493846	1.429103
Επ. σημαντικότητας	Ολα τα επίπεδα	Ολα τα επίπεδα

Πίνακας 16.8

```
===    Επιτυχίες    Αποτυχίες    Στατιστική Ανάλυση    Εξοδος
                                  Συνολικά στατιστικά στοιχεία
                                  Διαχωριστική ανάλυση
                                  Πίνακας ταξινόμησης
                                  ξινόμησης
                                  κασίας
                Ισότητα μέσων τιμών
                Ισότητα διασπορών
                Συντελεστης συσχέτισης
                Κανονικότητα της Z
                Παλινδρόμηση

Η εξίσωση παλινδρόμησης στις επιτυχίες έχει :
 a = -0.5739      b = -1.9639      R = -0.9900
 Αφού έγινε ο έλεγχος κατά Student προέκυψε
 για τον πληθυσμό R = 0.

   Πάτησε οποιοδήποτε πλήκτρο για συνέχεια.
```

Πίνακας 16.9

ΜΕΤΑΒΛΗΤΕΣ	ΜΕΣΗ ΤΙΜΗ	ΤΥΠ.ΑΠΟΚ/ΣΗ	ΜΕΓΙΣΤΟ	ΕΛΑΧΙΣΤΟ
DIAMETER..	23.734	20.108	255.800	3.000
DIM1.....	41.773	24.566	230.000	14.353
DIM2.....	48.416	57.169	427.160	0.328
DIM3.....	30.589	12.921	83.333	16.778
DIM4.....	13.574	7.812	34.668	0.484
DIM5.....	6.281	11.657	56.121	0.000

Πίνακας 17.1

ΜΕΤΑΒΛΗΤΕΣ	Ui	Si	ΚΑΤΑΤΑΞΗ	ΣΤΑΘΕΡΟΤΗΤΑ	Uo
DIAMETER..	0.003510	0.070582	5	ΑΣΤΑΘΗΣ	0.185508
DIM1.....	0.002950	0.072464	4	ΕΥΣΤΑΘΗΣ	
DIM2.....	0.013044	0.745689	1	ΕΥΣΤΑΘΗΣ	
DIM3.....	0.003622	0.046800	6	ΕΥΣΤΑΘΗΣ	
DIM4.....	-0.092390	-0.721734	2	ΕΥΣΤΑΘΗΣ	
DIM5.....	0.019065	0.222231	3	ΕΥΣΤΑΘΗΣ	

Πίνακας 17.2

Π Ι Ν Α Κ Α Σ Κ Α Ι Π Ο Σ Ο Σ Τ Α Τ Α Ξ Ι Ν Ο Μ Η Σ Η Σ

Ποσοστά Σωστής Ταξινόμησης

ΕΠΙΤΥΧΙΕΣ .... EF1 = 76.1%
ΑΠΟΤΥΧΙΕΣ .... EF2 = 86.4%
ΣΥΝΟΛΙΚΑ ..... EFt = 79.8%

		ΠΡΟΒΛΕΠΟΜΕΝΑ: ΕΠΙΤΥΧΙΕΣ	ΠΡΟΒΛΕΠΟΜΕΝΑ: ΑΠΟΤΥΧΙΕΣ	
ΠΡΑΓΜΑΤΙΚΑ	ΕΠΙΤΥΧΙΕΣ	83	26	109
	ΑΠΟΤΥΧΙΕΣ	8	51	59
		91	77	

ESCAPE : Επιστροφή

Πίνακας 17.3

=== Επιτυχίες Αποτυχίες Στατιστική Ανάλυση Εξοδος

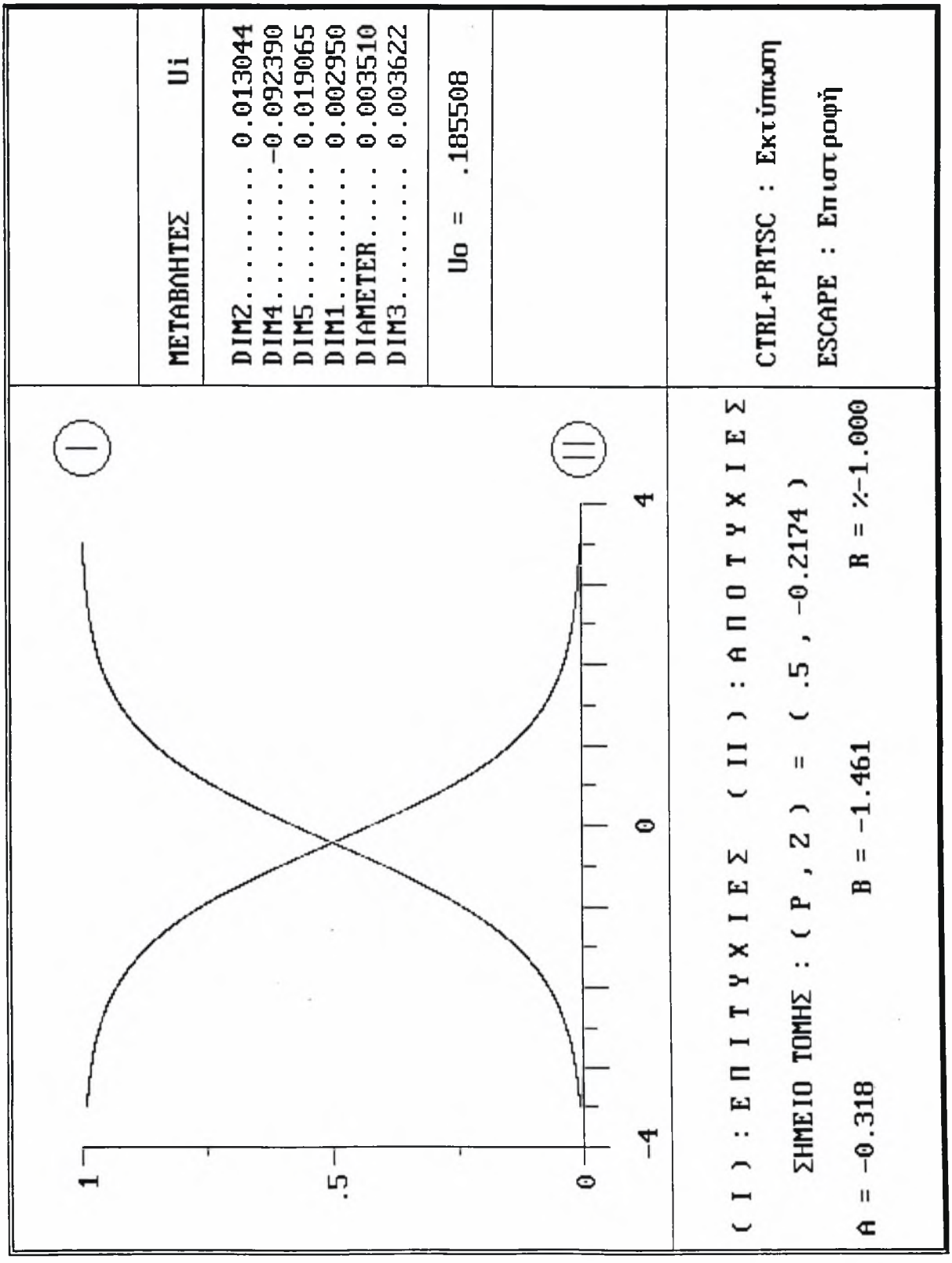
Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
ης Z

R A O ' S F			
Wilk's Λ	F	v1	v2
0.6701	13.2087	6	161

Πάτησε οποιοδήποτε πλήκτρο για επιστροφή

Πίνακας 17.4



Πίνακας 17.5

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
ης Z

B O X ' S F			
Box's B	F	v1	v2
418.4545	19.0540	21	53659
Πάτησε οποιοδήποτε πλήκτρο για επιστροφή			

Πίνακας 17.6

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

σων τιμών
ασπορών
ς συσχέτισης
τα της Z
ση

Ιδιοτιμές	C.C.C
0.00000	0.0000
0.49225	0.5743
0.00000	0.0004
0.00000	0.0001
-0.00000	0.0000
0.00000	0.0003
Πάτησε οποιοδήποτε πλήκτρο	

Πίνακας 17.7

≡≡≡ Επιτυχίες Αποτυχίες Στατιστική Ανάλυση Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών

Ε λ ε γ χ ο ς	K o l m o g o r o v - S m i r n o v	
	Ε π ι τ υ χ ί ε ς	Α π ο τ υ χ ί ε ς
Μέση τιμή	.5131025	-.947935
Μεταβλητότητα	1.3263	.3632779
Επ. σημαντικότητας	Ολα τα επίπεδα	Ολα τα επίπεδα

Πίνακας 17.8

≡≡≡ Επιτυχίες Αποτυχίες Στατιστική Ανάλυση Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
Κανονικότητα της Z
Παλινδρόμηση

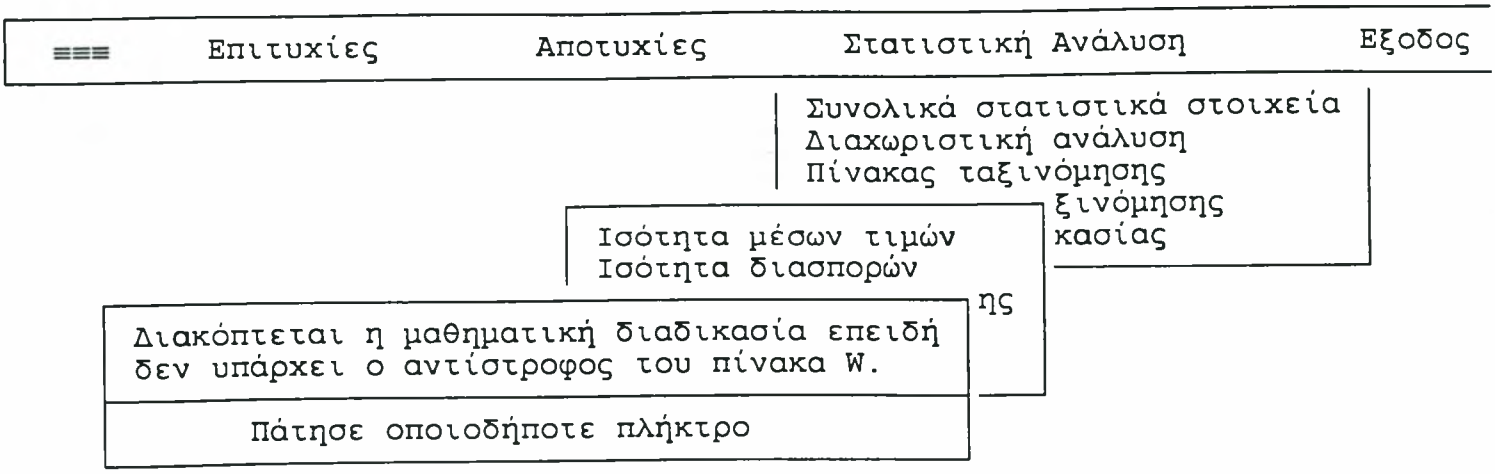
Η εξίσωση παλινδρόμησης στις επιτυχίες έχει :
a = -0.3177 b = -1.4610 R = -0.9900
Αφού έγινε ο έλεγχος κατά Student προέκυψε
για τον πληθυσμό R = 0.

Πάτησε οποιοδήποτε πλήκτρο για συνέχεια.

Πίνακας 17.9

ΜΕΤΑΒΛΗΤΕΣ	ΜΕΣΗ ΤΙΜΗ	ΤΥΠ.ΑΠΟΚ/ΣΗ	ΜΕΓΙΣΤΟ	ΕΛΑΧΙΣΤΟ
THICKNESS.	0.611	0.166	1.000	0.079
DIM1.....	41.773	24.566	230.000	14.353
DIM2.....	48.416	57.169	427.160	0.328
DIM3.....	30.589	12.921	83.333	16.778
DIM4.....	13.574	7.812	34.668	0.484
DIM5.....	6.281	11.657	56.121	0.000

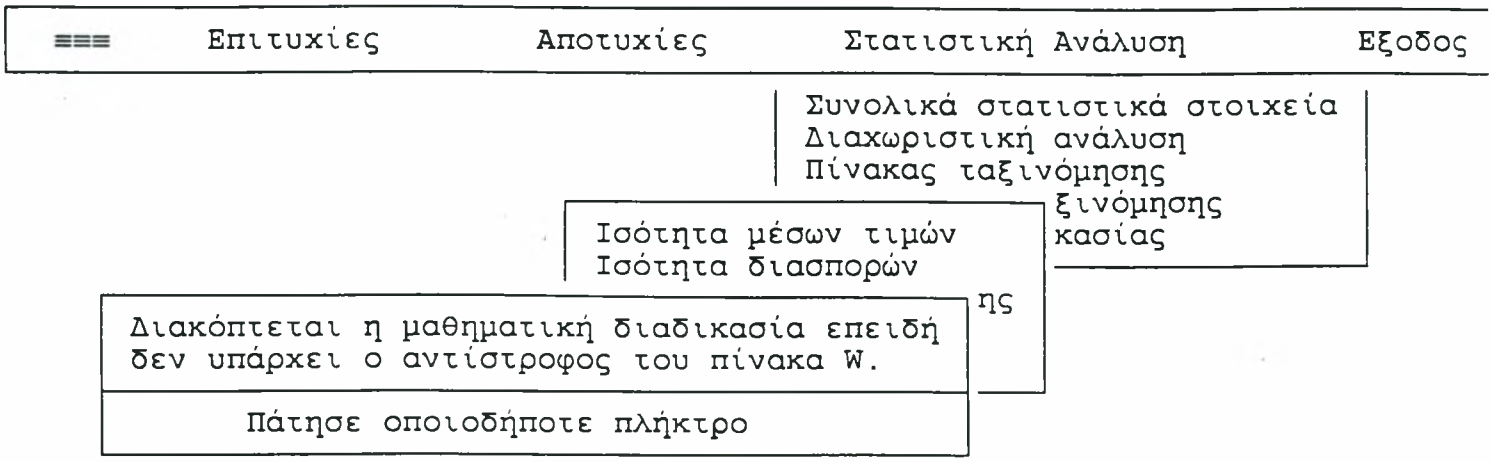
Πίνακας 18.1



Πίνακας 18.2

ΜΕΤΑΒΛΗΤΕΣ	ΜΕΣΗ ΤΙΜΗ	ΤΥΠ.ΑΠΟΚ/ΣΗ	ΜΕΓΙΣΤΟ	ΕΛΑΧΙΣΤΟ
PRESSURE..	112.633	31.498	255.000	18.000
DIM1.....	41.773	24.566	230.000	14.353
DIM2.....	48.416	57.169	427.160	0.328
DIM3.....	30.589	12.921	83.333	16.778
DIM4.....	13.574	7.812	34.668	0.484
DIM5.....	6.281	11.657	56.121	0.000

Πίνακας 19.1



Πίνακας 19.2

ΜΕΤΑΒΛΗΤΕΣ	ΜΕΣΗ ΤΙΜΗ	ΤΥΠ.ΑΠΟΚ/ΣΗ	ΜΕΓΙΣΤΟ	ΕΛΑΧΙΣΤΟ
LIFETIME..	64.720	37.543	221.000	0.000
DIM1.....	41.773	24.566	230.000	14.353
DIM2.....	48.416	57.169	427.160	0.328
DIM3.....	30.589	12.921	83.333	16.778
DIM4.....	13.574	7.812	34.668	0.484
DIM5.....	6.281	11.657	56.121	0.000

Πίνακας 20.1

ΜΕΤΑΒΛΗΤΕΣ	Ui	Si	ΚΑΤΑΤΑΞΗ	ΣΤΑΘΕΡΟΤΗΤΑ	Uo
LIFETIME..	-0.029855	-1.120839	1	ΕΥΣΤΑΘΗΣ	1.005167
DIM1.....	0.011547	0.283664	5	ΑΣΤΑΘΗΣ	
DIM2.....	-0.013183	-0.753664	2	ΕΥΣΤΑΘΗΣ	
DIM3.....	-0.002152	-0.027805	6	ΕΥΣΤΑΘΗΣ	
DIM4.....	0.061964	0.484056	4	ΕΥΣΤΑΘΗΣ	
DIM5.....	0.048984	0.570991	3	ΑΣΤΑΘΗΣ	

Πίνακας 20.2

Π Ι Ν Α Κ Α Σ Κ Α Ι Π Ο Σ Ο Σ Τ Α Τ Α Ξ Ι Ν Ο Μ Η Σ Η Σ

Ποσοστά Σωστής Ταξινόμησης

ΕΠΙΤΥΧΙΕΣ .... EF1 = 91.7%
ΑΠΟΤΥΧΙΕΣ .... EF2 = 98.3%
ΣΥΝΟΛΙΚΑ ..... EFt = 94.0%

		ΠΡΟΒΛΕΠΟΜΕΝΑ: ΕΠΙΤΥΧΙΕΣ	ΠΡΟΒΛΕΠΟΜΕΝΑ: ΑΠΟΤΥΧΙΕΣ	
ΠΡΑΓΜΑΤΙΚΑ	ΕΠΙΤΥΧΙΕΣ	100	9	109
	ΑΠΟΤΥΧΙΕΣ	1	58	59
		101	67	

ESCAPE : Επιστροφή

Πίνακας 20.3

=== Επιτυχίες Αποτυχίες Στατιστική Ανάλυση Εξοδος

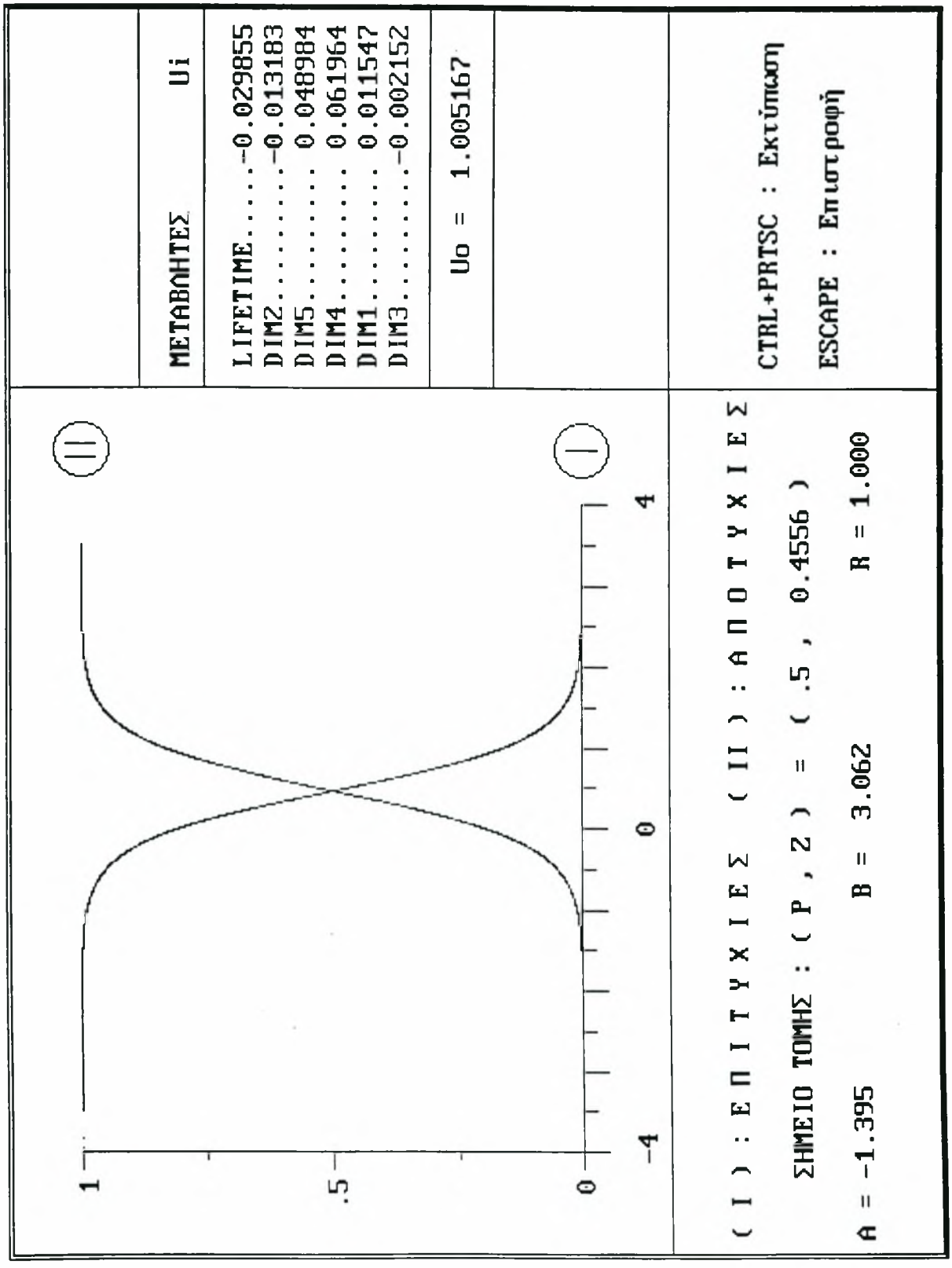
Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
ης Z

R A O ' S F			
Wilk's Λ	F	v1	v2
0.3163	57.9943	6	161

Πάτησε οποιοδήποτε πλήκτρο για επιστροφή

Πίνακας 20.4



Πίνακας 20.5

====	Επιτυχίες	Αποτυχίες	Στατιστική Ανάλυση	Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
ης Z

B O X ' S F			
Box's B	F	v1	v2
362.2751	16.4959	21	53659
Πάτησε οποιοδήποτε πλήκτρο για επιστροφή			

Πίνακας 20.6

====	Επιτυχίες	Αποτυχίες	Στατιστική Ανάλυση	Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

σων τιμών
ασπορών
ς συσχέτισης
τα της Z
ση

Ιδιοτιμές	C.C.C
0.00000	0.0000
0.00000	0.0000
0.00000	0.0005
0.00000	0.0003
2.16128	0.8268
-0.00000	0.0000
Πάτησε οποιοδήποτε πλήκτρο	

Πίνακας 20.7

≡≡≡	Επιτυχίες	Αποτυχίες	Στατιστική Ανάλυση	Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών

Ε λ ε γ χ ο ς	K o l m o g o r o v - S m i r n o v	
	Ε π ι τ υ χ ί ε ς	Α π ο τ υ χ ί ε ς
Μέση τιμή	-1.075146	1.986288
Μεταβλητότητα	1.274526	.4589253
Επ. σημαντικότητας	Ολα τα επίπεδα	Ολα τα επίπεδα

Πίνακας 20.8

≡≡≡	Επιτυχίες	Αποτυχίες	Στατιστική Ανάλυση	Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
Κανονικότητα της Z
Παλινδρόμηση

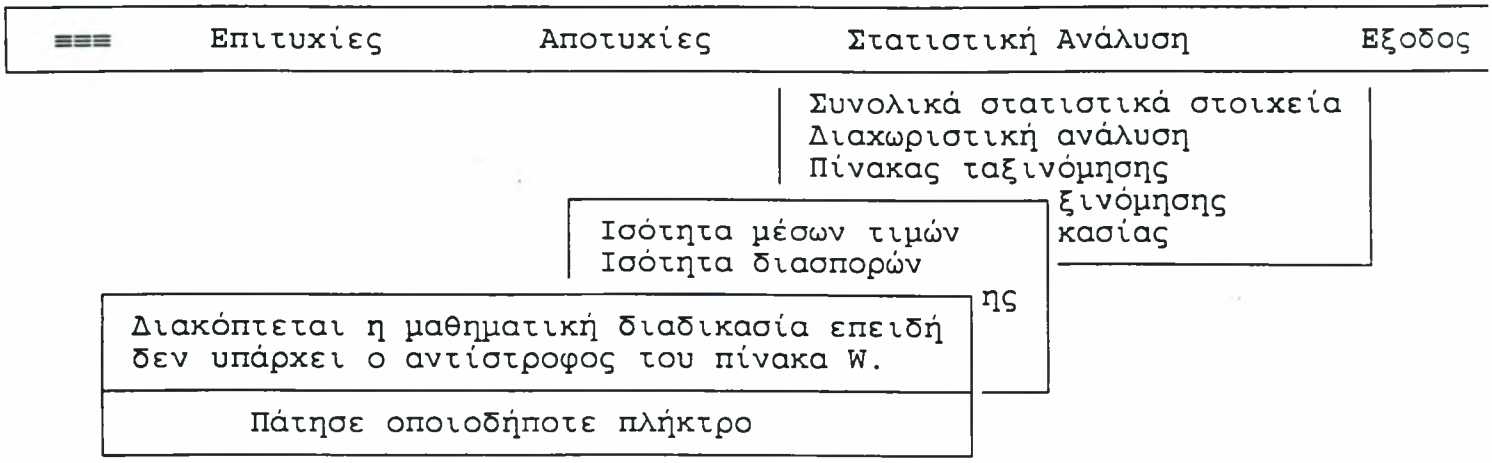
Η εξίσωση παλινδρόμησης στις επιτυχίες έχει :
a = -1.3948 b = 3.0616 R = 1.0000
Αφού έγινε ο έλεγχος κατά Student προέκυψε
για τον πληθυσμό R > 0.

Πάτησε οποιοδήποτε πλήκτρο για συνέχεια.

Πίνακας 20.9

ΜΕΤΑΒΛΗΤΕΣ	ΜΕΣΗ ΤΙΜΗ	ΤΥΠ.ΑΠΟΚ/ΣΗ	ΜΕΓΙΣΤΟ	ΕΛΑΧΙΣΤΟ
THICKNESS.	0.611	0.166	1.000	0.079
LIFETIME..	64.720	37.543	221.000	0.000
DIM1.....	41.773	24.566	230.000	14.353
DIM2.....	48.416	57.169	427.160	0.328

Πίνακας 21.1



Πίνακας 21.2

ΜΕΤΑΒΛΗΤΕΣ	ΜΕΣΗ ΤΙΜΗ	ΤΥΠ.ΑΠΟΚ/ΣΗ	ΜΕΓΙΣΤΟ	ΕΛΑΧΙΣΤΟ
LENGTH....	50.358	63.897	452.000	1.000
PRESSURE..	112.633	31.498	255.000	18.000
LIFETIME..	64.720	37.543	221.000	0.000
DIM3.....	30.589	12.921	83.333	16.778

Πίνακας 22.1

ΜΕΤΑΒΛΗΤΕΣ	Ui	Si	ΚΑΤΑΤΑΞΗ	ΣΤΑΘΕΡΟΤΗΤΑ	Uo
LENGTH....	-0.002044	-0.130607	4	ΕΥΣΤΑΘΗΣ	-2.020508
PRESSURE..	-0.005929	-0.186749	3	ΕΥΣΤΑΘΗΣ	
LIFETIME..	0.026040	0.977623	1	ΕΥΣΤΑΘΗΣ	
DIM3.....	0.036153	0.467137	2	ΕΥΣΤΑΘΗΣ	

Πίνακας 22.2

Π Ι Ν Α Κ Α Σ Κ Α Ι Π Ο Σ Ο Σ Τ Α Τ Α Ξ Ι Ν Ο Μ Η Σ Η Σ

Ποσοστά Σωστής Ταξινόμησης

ΕΠΙΤΥΧΙΕΣ .... EF1 = 86.2%
ΑΠΟΤΥΧΙΕΣ .... EF2 = 94.9%
ΣΥΝΟΛΙΚΑ ..... EFt = 89.3%

		ΠΡΟΒΛΕΠΟΜΕΝΑ: ΕΠΙΤΥΧΙΕΣ	ΠΡΟΒΛΕΠΟΜΕΝΑ: ΑΠΟΤΥΧΙΕΣ	
ΠΡΑΓΜΑΤΙΚΑ	ΕΠΙΤΥΧΙΕΣ	94	15	109
	ΑΠΟΤΥΧΙΕΣ	3	56	59
		97	71	

ESCAPE : Επιστροφή

Πίνακας 22.3

=== Επιτυχίες Αποτυχίες Στατιστική Ανάλυση Εξοδος

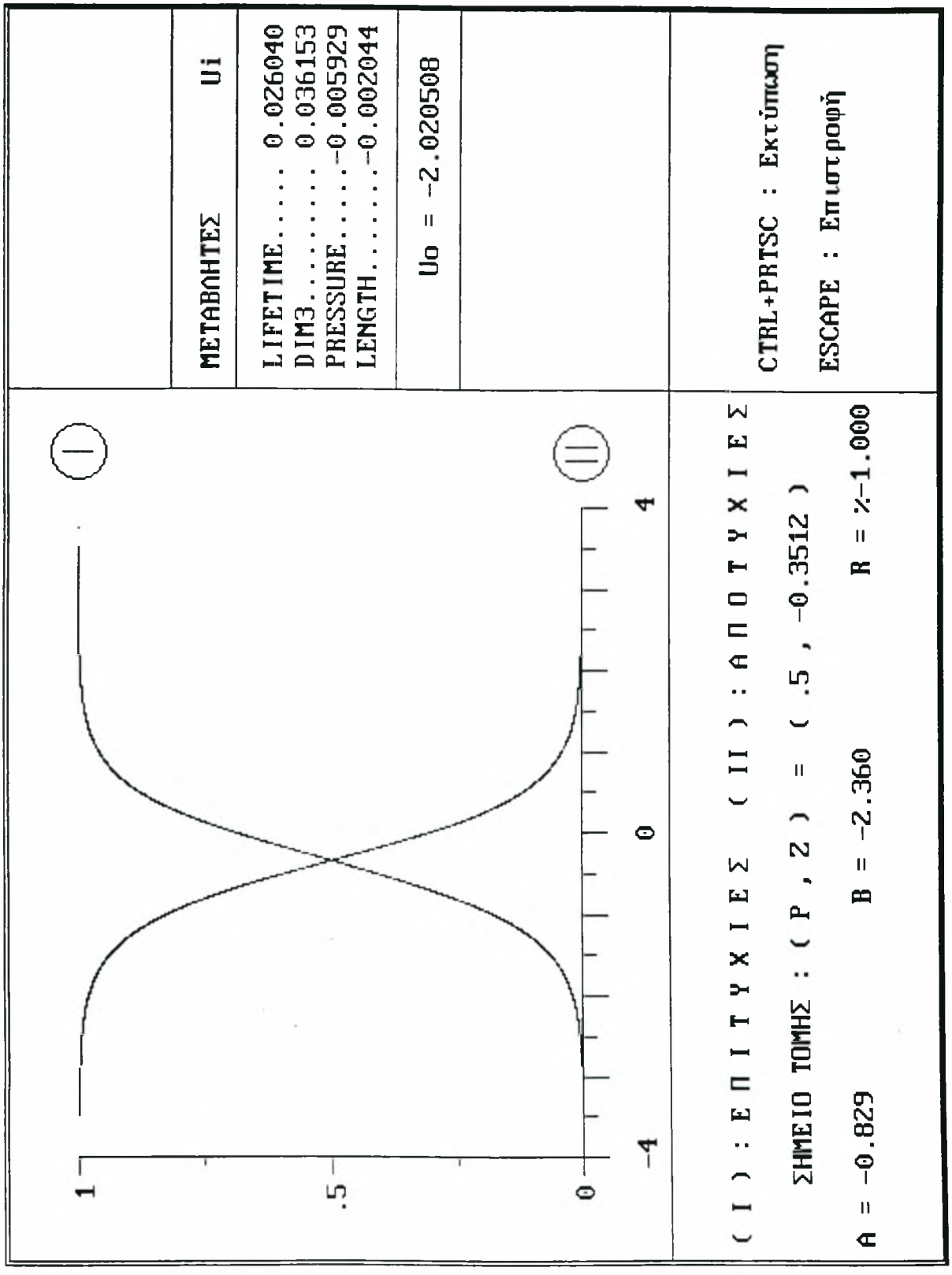
Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
ης Z

R A O ' S F			
Wilk's Λ	F	v1	v2
0.4377	52.3467	4	163

Πάτησε οποιοδήποτε πλήκτρο για επιστροφή

Πίνακας 22.4



Πίνακας 22.5

≡≡≡	Επιτυχίες	Αποτυχίες	Στατιστική Ανάλυση	Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
ης Z

B O X ' S F			
Box's B	F	v1	v2
115.3890	11.1985	10	67143
Πάτησε οποιοδήποτε πλήκτρο για επιστροφή			

Πίνακας 22.6

≡≡≡	Επιτυχίες	Αποτυχίες	Στατιστική Ανάλυση	Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

σων τιμών
ασπορών
ς συσχέτισης
τα της Z
ση

Ιδιοτιμές	C.C.C
0.00000	0.0000
1.28458	0.7499
-0.00000	0.0000
0.00000	0.0001
Πάτησε οποιοδήποτε πλήκτρο	

Πίνακας 22.7

≡≡≡	Επιτυχίες	Αποτυχίες	Στατιστική Ανάλυση	Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών

Ε λ ε γ χ ο ς	K o l m o g o r o v - S m i r n o v	
	Ε π ι τ υ χ ί ε ς	Α π ο τ υ χ ί ε ς
Μέση τιμή	.8288823	-1.531325
Μεταβλητότητα	1.235846	.5303861
Επ. σημαντικότητας	Κανένα επίπεδο	Κανένα επίπεδο

Πίνακας 22.8

≡≡≡	Επιτυχίες	Αποτυχίες	Στατιστική Ανάλυση	Εξοδος

Συνολικά στατιστικά στοιχεία
Διαχωριστική ανάλυση
Πίνακας ταξινόμησης
ξινόμησης
κασίας

Ισότητα μέσων τιμών
Ισότητα διασπορών
Συντελεστης συσχέτισης
Κανονικότητα της Z
Παλινδρόμηση

Η εξίσωση παλινδρόμησης στις επιτυχίες έχει :
a = -0.8290 b = -2.3602 R = -0.9900
Αφού έγινε ο έλεγχος κατά Student προέκυψε
για τον πληθυσμό R = 0.

Πάτησε οποιοδήποτε πλήκτρο για συνέχεια.

Πίνακας 22.9